3

REPORT RESUMES

ED 003 887 48

EXPERIMENTAL ANALYSIS OF THE CONTROL OF SPEECH PRODUCTION AND PERCEPTION-VII. FINAL REPORT.

BY- LANE, HARLAN

MICHIGAN UNIV., ANN ARBOR,COLL.OF LIT.,SCI.,ARTS

REPORT NUMBER NDEA-VI-78-7 PUB DATE FEB 65

REPORT NUMBER REPORT-7

REPORT NUMBER 05586-05613-4-F

CONTRACT OEC-3-14-013

EDRS PRICE MF-$0.09 HC-$2.32 58P.

DESCRIPTORS- *SPEECH, *ORAL COMMUNICATION, *LINGUISTICS, *LANGUAGE PATTERNS, EXPERIMENTAL PROGRAMS, BEHAVIOR, MODELS, SPEECH EDUCATION, *PERCEPTION, RETENTION, ANN ARBOR, MICHIGAN

THIS REPORT CONTAINS FOUR ARTICLES REPRESENTING THE FINAL EFFORTS OF A COMPREHENSIVE EXPERIMENTAL PROGRAM IN SPEECH CONTROL. THE ARTICLES ARE TITLED (1) "THE BEHAVIORAL BASIS OF THE POLARITY PRINCIPLE IN LINGUISTICS," (2) "MODELS OF LEARNING AND METHODS OF TEACHING," (3) "RETENTION OF VOCAL TOPOGRAPHY," AND (4) "MULTIDIMENSIONAL STIMULUS CONTROL OF THE DISCRIMINATIVE RESPONSE IN EXPERIMENTAL CONDITIONING AND PSYCHOPHYSICS." (RELATED REPORTS ARE ED 003 883-ED 003 887.) (JH)

ED003887

78
No 7

05586,05613-4-F

THE UNIVERSITY OF MICHIGAN

COLLEGE OF LITERATURE, SCIENCE, AND THE ARTS
DEPARTMENT OF PSYCHOLOGY
BEHAVIOR ANALYSIS LABORATORY

Final Report

Experimental Analysis of the Control of Speech Production and Perception: VII

FR (No. 7 in Series)

Supported by:

Language Development Section
U. S. Office of Education
Contract No. OE-3-14-013
Washington, D. C.

and

National Science Foundation
Grant No. GB-552
Washington, D. C.

Administered through: *February 1965*

OFFICE OF RESEARCH ADMINISTRATION • ANN ARBOR

U. S. DEPARTMENT OF HEALTH, EDUCATION AND WELFARE
Office of Education
This document has been reproduced exactly as received from the person or organization originating it. Points of view or opinions stated do not necessarily represent official Office of Education position or policy.

THE UNIVERSITY OF MICHIGAN

COLLEGE OF LITERATURE, SCIENCE, AND THE ARTS

Department of Psychology
Behavior Analysis Laboratory

FINAL REPORT

EXPERIMENTAL ANALYSIS
OF
THE CONTROL
OF
SPEECH PRODUCTION AND PERCEPTION: VII

Project Director Harlan Lane
Harlan L. Lane
Associate Professor of Psychology
The University of Michigan

supported by:

LANGUAGE DEVELOPMENT SECTION
U. S. OFFICE OF EDUCATION
CONTRACT NO. OE-3-14-013
WASHINGTON, D. C.

and

NATIONAL SCIENCE FOUNDATION
GRANT NO. GB-552
WASHINGTON, D.C.

administered through:

OFFICE OF RESEARCH ADMINISTRATION ANN ARBOR

February 1965

PRECEDING PAGE BLANK-NOT FILMED

TABLE OF CONTENTS

STAFF OF THE BEHAVIOR ANALYSIS LABORATORY

THE ARTICLES

THE BEHAVIORAL BASIS OF THE POLARITY PRINCIPLE IN LINGUISTICS, by Harlan L. Lane

MODELS OF LEARNING AND METHODS OF TEACHING, by Harlan L. Lane

RETENTION OF VOCAL TOPOGRAPHY, by Harlan L. Lane, James L. Kopp and William L. Sheppard

MULTIDIMENSIONAL STIMULUS CONTROL OF THE DISCRIMINATIVE RESPONSE IN EXPERIMENTAL CONDITIONING AND PSYCHOPHYSICS, by D. V. Cross

STAFF MEMBERS

of the

BEHAVIOR ANALYSIS LABORATORY*

H. L. Lane, Ph.D., Director

S. A. Ross, Mag. Art., Research Associate

D. V. Cross, A.M., Research Assistant

J. L. Kopp, A.M., Research Assistant

R. L. Buiten, B.S.E., Research Engineer

D. R. Brinkman, B.S.E., Administrative Assistant

W. C. Sheppard, B.A., Research Assistant

R. E. Lehman, Technical Assistant

G. A. Lombard, Technical Assistant

Timothy C. Rand, Research Technician

L. G. Peterson, Secretary

---

* In February 1965 the Behavior Analysis Laboratory was assimilated by the University of Michigan Center for Research on Language and Language Behavior.

THE BEHAVIORAL BASIS OF THE POLARITY PRINCIPLE IN LINGUISTICS

Harlan L. Lane
The University of Michigan

The Behavioral Basis of the Polarity Principle in Linguistics[1]

Harlan L. Lane

Center for Research on Language and Language Behavior

The University of Michigan

'Phonemes do not differ as five-inch or six-inch entities differ, but as clubs and poles differ.' (Edward Sapir)

There are at least seven major linguists who have argued, in defense of the magic number two, that plus and minus--binary opposition--are the fundamentals of language. The premise that the dichotomous scale is inherent in the structure of language may be called the polarity principle of linguistics, thereby acknowledging its central and pervasive role in the phonological component of language. Roman Jakobson and Morris Halle, who introduced the phrase 'polarity principle', emphasize that it refers to more than a mere principle of analysis in linguistic studies:

> 'A system of distinctive features based on a mutually implicating relation between the terms of each binary opposition is the optimal (linguistic) code, and it is unwarranted to assume that the speech participants in their encoding and decoding operations use a more complicated and less economic set of differential criteria...the inherence of the dichotomous scale in the linguistic system is quite manifest.'[2]

This article examines the ways in which the polarity principle inheres in the decoding and encoding operations of language users. Statements about the polarity principle imply descriptions of the following two behaviors: detecting that two speech sounds are different and identifying each of them. The implied descriptions are spelled out first. Then, evidence is presented that the polarity principle leads in fact to an accurate description of the discrimination and identification of speech sounds. Finally, the article explains the basis of the findings that identification and discrimination are related to each other and to the acoustic continuum in the peculiar ways provided by the polarity principle. In so doing, it casts doubt upon a rather popular explanation that attributes the properties of identification and discrimination to articulation. There is reason to believe that the descriptions of language behavior provided by the polarity principle are, under certain conditions, equally valid as descriptions of any behavior, human or infrahuman--in other words, that the polarity principle is merely the formulation in the domain of language of some very general laws concerning the behavior of organisms.

Behavioral implications of the polarity principle

The polarity principle is not only the basic tenet of the theory of distinctive features, its touchstone in the search for linguistic universals, it is, more generally, an axiom of most--if not all--approaches to phonology. Charles Hockett uses the principle to delimit the boundaries of language:

> 'No phonologically relevant contrasts are of the continuous-scale type. In a language where voicing and voicelessness are phonologically functional, a given bit of speech is either voiced or voiceless--not, structurally half-voiced or three-quarters voiced...If we find continuous scale contrasts in

> the vicinity of what we are sure is language, we exclude them from language...'[3]

It is Sapir, however, who has most clearly stated the implications of the polarity principle for a description of the behavior of the language user:

> 'Phonemes do not differ as five-inch or six-inch entities differ but as clubs and poles differ. If the phonetician discovers in the flow of actual speech something that is neither "club" nor "pole", he, as phonetician, has the right to set up a "half-way between club and pole" entity. Functionally, however, such an entity is a fiction and the naive speaker or hearer is not only driven by its relational behavior to classify it as a "club" or a "pole," but actually hears and feels it to be such.'[4]

Clearly, the polarity principle entails a specific description of the way in which identification responses are distributed over an acoustic-phonetic continuum. The behavior imposes a dichotomous scale on the continuum: one response is evoked, always with unit probability, over a range of stimulus values; then there is a 'phoneme boundary' after which another response is evoked with unit probability over another range of stimulus values. Figure 1 shows the graphic realization of the polarity principle as it is applied to the identification of speech sounds. C. L. Ebeling is quite explicit about the behavioral

----

Insert Fig. 1 about here

----

mapping implied by the principle:

> 'A sound dimension running from A to Z when parcelled out phonemically consists of...A - P allotted to a phoneme, a no-man's-land P - Q, and Q - Z allotted to another phoneme...the function of a point is wholly defined by the mere fact that it belongs to one of the segments.'[5]

Figure 1 shows, then, the theoretical description of the identification of speech sounds provided by the polarity principle. No continuous scale, but rather a partitioning of the acoustic-phonetic continuum, is entailed.

Now consider the implications of the principle for the behavior of detecting a difference between two speech sounds. When the description of language behavior implied by the polarity principle is made explicit, it becomes clear that the principle underlies that most basic of operations in phonology, the Pair Test. The Pair Test is a procedure for classing together these physically different utterances that are not distinguished by a listener and, conversely, for separating those physically different utterances that are discriminated. Noam Chomsky calls the Pair Test "one of the operational cornerstones for linguistic theory."[6] Without the Pair Test, or some related way of making contact with the behavior of the listener, there could be no phonemics, for it is the Pair Test that enables a virtually limitless number of sounds to be grouped into functional classes. Hockett describes the basis of the test in this way:

> 'The fundamental assumption in phonology is that, despite the wide variation of actual articulatory motion from one speech event to another in a community, the speakers of a language themselves hear

some speech events as the same, some as different.'[7]

The Pair Test presupposes that phonetically distinct utterances heard as 'the same' come from the same phoneme class whereas those that are heard as 'different' come from different phoneme classes. The phoneitc distinction is labeled free variation in the first case, and it is labeled contrast in the latter case. Figure 2 shows graphically the way in which the Pair Test presupposes the polarity principle.

Insert Fig. 2 about here

At the top of the figure is shown the distribution of same-different or discriminative responses and at the bottom the distribution of identification responses of our hypothetical listener. Pairwise comparisons of points within the first phoneme class lead to judgments of 'same' (zero percent 'different'). A comparison of two physical values drawn from diffferent phoneme classes leads to judgments of 'different.' Finally, pairwise comparison of phonemes in free-variation within the second phoneme class lead again to judgments of 'same.'

Just as the polarity principle excluded continuous-scale contrasts in the realm of identification, so, too, the Pair Test presupposes the exclusion of varying degrees of discrimination. In the absence of the operation of the polarity principle, however, varying degrees of discrimination would often be obtained. Psychophysical studies tell us that a fixed increment in stimulus value is increasingly difficult to detect with increasing stimulus magnitude and that it takes very small increments indeed to cause the listener to fail to detect a difference all of the time. How would the phonemic analysis proceed if the Pair Test kept turning up gradual discrimination functions--if the listener called two sounds the same, say 70 percent of the time, but different 30 percent of the time? Of course, the analysis could not proceed if discriminations among speech sounds were of the 'continuous scale' type. Because of the operation of the polarity principle, however, the Pair Test is safe in assuming, within error of measurement, an underlying dichotomous scale: free variation within phoneme classes, contrast between them.

Finally, Figure 3 shows some implications drawn rather indirectly from the polarity principle concerning the distribution of two more behavioral measures: the latency and the topography of identification responses. With regard to the latency of the identification responses, we might speculate that speech sounds drawn from within the phoneme classes would be identified as quickly as they are reliably . Alvin M. Liberman, of Haskins Laboratories, has put it this way:

Insert Fig. 3 about here

> 'Because of the discrimination peaks at the phoneme boundaries, the incoming sounds are heard categorically... and they are therefore quickly and accurately sorted into the appropriate phoneme bins.'[8]

On the other hand, speech sounds drawn from the immediate region of no-man's-land--that is, from the boundary between phoneme classes--should lead to response equivocation and to increased latency of responding.

With regard to the topography or phonetic properties of the identification responses, the function shown is attributed to the polarity principle in view of Jakobson and Halle's third argument for the inherence of the binary scale in the structure of language:

> 'Almost all of the distinctive features show an unquestionably dichotomous structure on their acoustical and, correspondingly, on their motor level.'[9]

It is, of course, possible that responses to speech sounds drawn from the region of the phoneme boundary would show some measure of response 'blending'. For example, it is conceivable that an interovocalic plosive with a duration intermediate to those for /p/ and /b/ would evoke an identification response with some intermediate topography; this is possible, but it seems most consonant with the polarity principle to predict that either /p/ or /b/ will be emitted and nothing in between, as shown in Figure 3.

Properties of identification and discrimination

Experiments with synthesized speech make it possible to assess the validity of the descriptions of identification and discrimination derived from the polarity principle.

Figure 4 shows the way in which a representative listener identified the stimuli in a series of synthetic speech sounds sampled from an acoustic continuum. The continuum in this study, conducted by Liberman and others,[10] ranged phonemically from /do/ to /to/ but the finding of a dichotomous behavioral scale has been replicated many times with a wide variety of phonemic contrasts. The data for individual listeners may differ in the precise point at which the phoneme boundaries are located but they agree in showing a boundary; continuous-scale contrasts

----
Insert Fig. 4 about here
----

are not in evidence.

The description of discrimination based on the polarity principle is also borne out by experimental findings. When the discriminability of pairs of stimuli from the /do/-/to/ continuum was determined for the subject whose identification functions are shown in Figure 4, the discrimination function plotted in Figure 5 was obtained. The method employed was the familiar ABX technique, in which failure of discrimination appears as 50 percent, or chance, accuracy. The hypothesized pattern of failure of discrimination within phoneme classes and enhancement of discrimination between phoneme classes is found in Figure 5, as it is whenever discrimination is measured along a continuum for which the polarity principle is operative. The intimate relation between the discrimination and identification functions is emphasized by the plot in Figure 5 of an 'expected' discrimination function. The expected founction is derived algebraically from the identification function on the simple assumption that two speech stimuli are discriminated only to the extent that they are assigned to different phoneme classes.

Insert Fig. 5 about here

Figure 6 shows that the latency of the identification responses is greatest when the listener tries to identify stimuli drawn from the region of the phoneme boundary and least when he is identifying speech sounds drawn from within phoneme classes. Once again, the data shown are illustrative of the findings of many experiments with diverse phonemic

Insert Fig. 6 about here

contrasts.

The case in support of the polarity principle seems equally clear with regard to the observed topography of identification responses. Katherine Harris, Jarvis Bastian and Alvin Liberman[11] have shown that when subjects were instructed to mimic stimuli sampled from an acoustic-phonetic continuum that ranged from /slIt/ to /splIt/, the silent intervals in their responses fell into a bimodal frequency distribution. To determine whether there might not be incipient lip movements evoked by stimuli in the region of the phoneme boundary, electromyographic recordings were made during identification; these were also entirely dichotomous. In the author's words: 'Either there was a normal burst of muscle potential at the lips (indicating a /p/ gesture) or there was not.'

The genesis of polarity

To recapitulate, there is a polarity principle at work in the decoding of speech sounds which is manifested in the probablilty, latency and topography of identification responses and in the accuracy of discriminative responses. A recent theory postulates that the source of the relations among these behavioral measures and, therefore, the underpinnings of the polarity principle are to be found in the articulation of speech sounds. This motor theory of speech perception is best characterized by Liberman, one of its major proponents:

> 'We believe that in the course of his long experience with language a speaker (and listener) learns to connect speech sounds with their appropriate articulations. In time, these articulatory movements and their sensory feedback... become part of the perceiving process, mediating between the acoustic stimulus

> and its ultimate perception. When significant acoustic cues that occupy different positions along a single continuum are produced by essentially discontinuous articulations...the perception becomes discontinuous (i.e., categorical) and discrimination peaks develop at the phoneme boundary.'[12]

The skeletal structure of the motor theory is simply this: the decoding process is found to impose polarity on acoustic continua. Decoding is linked by long experience with the language to encoding, which is essentially polar. Therefore, decoding is accomplished by reference to encoding. To say that speech is best analyzed scientifically by reference to articulation is one thing; to say that it is perceived by reference to articulation is quite another matter. The latter point of view, although popular, is grievously lacking in empirical support.[13] Perhaps the most cogent criticism of the motor theory of speech perception is that a series of stimuli can be selected that is not generated by articulation or by any of the encoding behaviors of the subject, and the relations between discrimination and identification cited by the motor theory can then be obtained quickly in the laboratory after a few minutes of conditioning. Under these conditions it is implausible to explain the resultant manifestations of the polarity principle in terms of articulation. A more likely inference is that the conditioning procedures used in the laboratory to establish polarity were also used in some fashion in the language community. In other words, if it can be shown that the four manifestations of the polarity principle found in language behavior can also be obtained by a certain conditioning procedure applied to virtually any behavior, it seems reasonable to attribute polarity to the conditioning procedure.

The conditioning procedure that provides the behavioral basis of the polarity principle is quite simple. Two stimuli are selected from a physical continuum and two responses are selected, or differentiated, from the repertory of the subject. In the presence of the first stimulus, occurrences of the first response are reinforced, whereas occurences of the second response are not. In the presence of the second stimulus, the relations between responses and reinforcement are reversed. Thus, the first stimulus is S^D for R_1 and S^Δ for R_2, whereas the second stimulus is S^D for R_2 and S^Δ for R_1. It will be convenient to have a name for this procedure: I will refer to it here as the Cross-Lane paradigm, since David Cross and I have employed it in several experiments.[14] This paradigm provides that, given two discriminative stimuli and two responses, only the first response is reinforced in the presence of the first stimulus and only the second response is reinforced in the presence of the second stimulus.

Consider first the properties of stimulus identification when the Cross-Lane paradigm is carried out with an acoustic continuum and with human vocal responses reinforced by money.[15] The two discriminative stimuli were half-second bursts of noise at sound pressure levels of 56 and 74 decibels; they were selected from a continuum ranging from 50 to 80 db (SPL). The two vocal responses were differentiated out of the subjects' repertoires by selective reinforcement; they were humming at 147 cps and at 227 cps, plus or minus 2 cps. Each subject was given 140 conditioning trials in which responses at 147 cps to the 56-db stimulus and responses at 227 cps to the 74 db stimulus were reinforced; the converse occurences were never reinforced. The perceptual structure imposed on the intensity continuum by this conditioning procedure was determined next by presenting stimuli sampled over a broad range and noting the probability, latency and topography of the identification responses.

The results of this experiment are displayed in Figure 7, where the operation of the polarity principle is manifest. In the first place, the identification responses impose a dichotomous scale on the continuum of noise intensities. In approximately the same degree as noted for the phonemic contrast /do/-/to/ 'the function of a point is wholly defined by the mere fact that it belongs to one of the segments.' In the second place, the binary scale is found to characterize the topography of the identification responses. When a stimulus intensity was presented, one of two response pitches was emitted: approximately, 147 cps or 227 cps. There were no gradations in topography, despite the fact that the subject had the articulatory capability of producing intermediate pitches. In the third place, the latency function characteristically associated with the polarity principle is obtained: the latencies of identification responses are minimal when stimuli drawn from within 'phoneme' classes are presented and the latencies are maximal when stimuli drawn from no-man's-land, in the region of the 'phoneme' boundary, are presented.

Insert Fig. 7 about here

A second example of the operation of the polarity principle with nonlinguistic continua was provided when the Cross-Lane paradigm was applied to a series of complex tones made up of two resonance bands that varied in their relative onset times.[16] These complex tones were actually the control stimuli in the study of the /do/-/to/ contrast described earlier. The control stimuli were prepared in this way: the hand-painted spectrograms which, when converted into sound by Pattern Playback,[17] ranged from /do/ to /to/, were inverted and then converted into sound. This operation preserved the stimulus variable (the delay in onset of the first formant) but it rendered the series of sounds unintelligible. The response /do/ was then conditioned to the inverted /do/ stimulus at the one end of the continuum (to my ear it sounded like 'bwank') and the response /to/ was conditioned to the inverted /to/ stimulus at the other extreme of the continuum (which also sounded like 'bwank'), following the Cross-Lane paradigm. Three college students were employed; the reinforcer during discrimination training was money. The resulting identification functions are shown in Figure 8. Clearly, the dichotomous scale inheres in the perception of complex tones--once the Cross-Lane paradigm is applied. These identification functions are directly comparable to those obtained for the /do/-/to/ continuum by Liberman et al. (Fig. 4),[18] since the stimulus variable is the same. The comparison reveals that the partitioning of nonspeech continua can be as unequivocally binary as the partitioning of speech continua.

Insert Fig. 8 about here

The operation of the polarity principle, following the application of the Cross-Lane paradigm to the /do/-/to/ control continuum, was also revealed when we examined the discriminability of the control stimuli. Figure 9 shows how accurately the second subject discriminated between two stimuli separated by two steps or by three steps at various points along the /do/-/to/ control continuum. The relations among discrimination, identification and the acoustic continuum are essentially the same as those

----
Insert Fig. 9 about here
----

obtained with the comparable speech continuum (shown previously in Figure 5). As provided by the polarity principle, the distinct sounds from within one 'phoneme class' were perceived as the same, whereas those drawn from different 'phoneme classes' were perceived as different. The dotted curve is the 'expected' discrimination function derived from the identification function on the simple assumption that the two speech stimuli are discriminated only to the extent that they are assigned to different phoneme classes. Comparison of the expected and obtained discrimination functions reveals that the polarity principle almost completely governs the detection of differences between these nonspeech sounds: when the differences became gross enough, however, psychophysical cues enter into account. The latter reservation is true of all the speech and non speech continua that have been invested in this context. The operation of the polarity principle was also demonstrated by the identification and discrimination functions for the other two subjects who participated in this experiment. Before the three subjects were conditioned according to the Cross-Lane paradigm, however, they failed to discriminate any of the stimulus comparisons at better than chance levels of accuracy.

The sudden manifestation of the polarity principle following the execution of the Cross-Lane paradigm is also found in a study of the perception of the /do/-/to/ continuum by an aphasic adult.[19] As Figure 10 shows, the series of synthetic speech stimuli were all perceived as /do/ by this aphasic and no differences were reliably detected among them, prior to conditioning in our laboratory. Following seven minutes of conditioning with the Cross-Lane paradigm, there is some

----
Insert Fig. 10 about here
----

evidence for the perceptual organization of the continuum into two classes; polarization was not complete, however. The discrimination function also shows a marked change in a direction that accords with the polarity principle: stimuli whose phonemic assignments are most disparate are most readily distinguished.

Further evidence of the operation of the polarity principle with arbitrary stimulus continua is provided by an experiment in which the Cross-Lane paradigm was applied to visual patterns.[20] The stimuli were circles, ten inches in diameter, with deleted sectors ranging from 30° to 62° in steps of 4°. Two discriminative responses were conditioned to the two extreme stimuli in the series and then intermediate stimuli were presented in a quasirandom order. Figure 11 shows the dichotomous identification function and the triangular latency function that we have come to associate with the polarity principle. The figure permits a detailed examination of the latencies of the identification responses of this subject, whose data are representative of those for 23 other college students who participated in this experiment. Most of the stimuli are reliably and quickly sorted into the appropriate 'phoneme' bins. Stimulus 4, however, falls precisely at the class boundary for this subject and he equivocates on the assignment of this stimulus, taking two tenths of a second on the average to decide on its assignment. Four other subjects in this experiment were conditioned in similar fashion with a reduced stimulus set, containing only stimuli 4 through 7. Figure 12 shows the characteristic dichotomous identification functions and the corresponding discrimi-

----
Insert Fig. 11 about here

Insert Fig. 12 about here
----

nation functions for each of the subjects. Once again, the application of the Cross-Lane conditioning paradigm leads to the operation of the polarity principle. The stimulus values selected from a continuum are perceived categorically. Stimuli are discriminated much more accurately if they are drawn from different categories than if they are drawn from the same category; witness the reliably better discrimination of stimuli 5 and 6 than stimuli 4 and 5, or 6 and 7.

Thus, the polarity principle will be found to operate whenever the Cross-Lane paradigm is applied: the operation of the principle in the realm of language behavior is a special case in several respects. First, the outcome of the paradigm is invariant with respect to sensory modality; the continuum need not be acoustic. Second, the product of the paradigm is invariant with respect to responses; these need not be verbal, nor articulatory, and they need not have in advance any unique relation to the stimulus set. Third, the Cross-Lane paradigm may be generalized so as to yield several partitionings of a stimulus continuum, rather than just one. For example, if three responses are selected and conditioned disjunctively to three discriminative stimuli located at various points on a continuum, three equivalence classes will be obtained. Within each class a set of stimuli are mutually substitutable in controlling a discrete response selected from a set of mutually exclusive and exhaustive responses. Between each pair of classes there is opposition or contrast. So it is that the behavioral hypotheses of the theory of distinctive features form a special case of the Cross-Lane paradigm in which (1) the stimuli are auditory; (2) the responses are linguistic; (3) the stimulus ranges are each partitioned only once; and (4) a particular set of controlling dimensions is postulated.

Let us explore a little further the third kind of invariance of the Cross-Lane paradigm, according to which several oppositions are contrasts may be set up along one continuum, so that the stimuli are substitutable within classes and exclusive between classes. This functional relation will be found in the identification responses conditioned by the language community to stimuli on the hue continuum. Since numerous investigators have reported that segments of the hue continuum are perceived categorically, it is tempting to speculate that something akin to the Cross-Lane paradigm is operative during the learning of color-naming and that adults will not only perceive colors categorically but also, in accord with the polarity principle, discriminate better between wave-lengths drawn from different color classes than between those drawn from within one color class. This speculation proves correct. Identification functions for spectral colors yield class boundaries at wave-lengths of 450, 490, 575 and 620 millimicrons.[21] According to the polarity principle, discrimination enhancement or, in other terms, the smallest values of the jnd should occur at these wave-lengths, and the findings of several investigators show that they do.[19] Moreover, the expected increase in the time required to identify stimuli drawn from the region of the class boundaries has also been demonstrated.[22]

There is reason to believe that a fourth kind of invariance should be attributed to the paradigm that we have been examining: this paradigm has been employed with aphasics, as illustrated earlier, with retardates, schizophrenics, rats and pigeons, and all have provided substantially the same results. Figure 13 shows the identification functions obtained from a severly retarded child.[23] This little girl was

----
Insert Fig. 13 about here
----

born with a congenital cerebral defect: she was totally blind and mute.

She was institutionalized when three years old and, nine years later, at the time of the experiment, her stanford-Binet I.Q. was 30. The Cross-Lane paradigm was applied with half-second bursts of tone at 56 and 74 db (SPL) and with two-button press responses; the reinforcer was 3 seconds of recorded children's music. When stimuli at various intensities were presented, ranging from 50 to 80 db (SPL), the characteristic binary scale was obtained.

The preceding experiment has been replicated with acute and with chronic schizophrenic patients.[24] Illustrative findings are plotted in Figure 14 for a 20-year-old college student with a diagnosis of acute schizophrenia. A comparable application of the Cross-Lane paradigm with noise intensities and key-peck responses in the pigeon led to similar results.[25] Finally, the paradigm has been applied with rats conditioned to emit one of two barpress responses to one of two light intensities.[26] Once again, categorical perception was demonstrated.

Insert Fig. 14 about here

Conclusion

Two responses may be differentiated in topography and then brought under discriminative control, either by the language community or by an experimenter following the Cross-Lane paradigm. If, as a consequence, two discriminative stimuli, drawn from a physical continuum, control two distinct responses and then stimuli intermediate on the continuum are presented, certain characteristic relations obtain among these stimuli and four properties of the behavior that they evoke--namely, the probability, latency and topography of identification responses and the accuracy of discrimination responses. These characteristic functional relations were first observed with linguistic continua and responses and given the collective name, in that context, of 'the polarity principle.' The relations are not, however, peculiar to language behavior but rather to the behavior of organisms, of which language behaviors are a small--albeit diverting--subset.

Footnotes

1. An address to the Center for Cognitive Studies, Harvard University, March 11, 1965. The Center for Research on Language and Language Behavior is supported in part by a contract with the Language Development Branch, U. S. Office of Education.

2. Roman Jakobson and Morris Halle, Fundamentals of language, 47-9 (The Hague, 1956).

3. Charles Hockett, A manual of phonology, 17 (Balitmore, 1955).

4. Edward Sapir, 'The psychological reality of the phoneme,' in D. G. Mandelbaum (Ed.), Selected writings of Edward Sapir, 47 (Berkeley, 1949).

5. C. L. Ebeling, Linguistic units, 45 (The Hague, 1962).

6. Noam Chomsky, 'The logical structure of linguistic theory,' unpublished manuscript (Cambridge, 1955).

7. Hockett, op. cit, 144.

'We say that the sounds occurring in complementary distribution are allophones of a phoneme, and not separate phonemes. In addition--and this is, to my mind, of the greatest weight--a native informant, if asked whether the word lull has or has not the same sound at the beginning and at the end, will invariably affirm that it has.' Ernest Pulgrain, Introduction to the spectrography of speech, 21 (The Hague, 1959).

8. Alvin M. Liberman, Franklin S. Cooper, Katherine S. Harris, and Peter F. MacNeilage, 'A motor theory of speech perception,' in C. G. M. Fant (Ed.), Proceedings of the Speech Communication Seminar (Stockholm, 1963).

9. Jakobson and Halle, op. cit., 48.

10. Alvin M. Liberman, Katherine S. Harris, Jo Ann Kinney, and Harlan L. Lane, 'The discrimination of relative onset time of the components of certain speech and nonspeech patterns,' J. exp. Psychol. 61.379-88 (1961).

11. Katherine S. Harris, Jarvis Bastian, and Alvin M. Liberman, 'Mimicry and the perception of a phonemic contrast induced by silent interval: electromyographic and acoustic measures,' _J. acoust. Soc. Amer._, 33.842 (1961).

12. Alvin M. Liberman, Katherine S. Harris, Peter Eimas, Leigh Lisker, and Jarvis Bastian, 'An effect of learning on speech perception: the discrimination of durations of silence with and without phonemic significance,' _Lang. and Speech_, 4.175-95 (1961).

13. Harlan L. Lane, 'The motor theory of speech perception: a critical review,' _Psychol. Rev._, in press (1965).

14. David V. Cross and Harlan L. Lane, 'On the discriminative control of concurrent responses: the relations among response frequency latency and topography in auditory generalization,' _J. exp. anal. Behav._, 5.487-96 (1962). Also see fn. 13, 16, 17, 20, 21, 22.

15. Ibid.

16. Lane, 'The motor theory...'

17. For a description of Pattern Playback and its principles of operation, see Franklin S. Cooper, 'Some instrumental aids to research on speech,' _Georgetown University Monogr. Lang. and Ling._ (Washington, 1953).

18. Liberman et al, 'An effect of learning...'

19. Harlan L. Lane, and Dennis J. Moore, 'Reconditioning a consonant discrimination in an aphasic: an experimental case history,' _J. Speech Hear. Disord._, 27.232-43 (1962).

20. David V. Cross, Harlan L. Lane, and William C. Sheppard, 'Identification and discrimination functions for visual continuum and their relation to the motor theory of speech perception,' _J. exp. Psychol._, in press (1965).

21. A. C. Beare, 'Color name as a function of wavelength,' *Amer. J. Psychol.*, 76.248-56 (1963).

22. L. A. Jones, 'The fundamental scale of pure hue and retinal sensibility to hue differences,' *J. opt Soc. Amer.*, 1.63-77 (1917). G. Ekman, 'Contributions to the psychophysics of color vision,' *Studium Generale*, 16.54-64 (1963).

23. Harlan L. Lane and Charles R. Curran, "Gradients of auditory generalization for blind retarded children,' *J. exp. Anal. Behav.*, 6.585-588 (1963).

24. Curran, 'Progress report on concurrent generalization in acute and chronic schizophrenics (with an appendix on concurrent generalization of two light intensities in rats),' unpublished manuscript, Department of Psychology, Columbia University (New York 1965).

25. David V. Carter, 'Progress report on concurrent generalization with pure-tone intensities and keypeck responses by the pigeon,' unpublished manuscript, Department of Psychology, Columbia University (New York, 1965).

26. Curran, 'Progress report...'

Figure Captions

Fig. 1. Hypothetical distribution over an acoustic-phonetic continuum of two identification responses, according to the polarity principle.

Fig. 2. Hypothetical distributions of discrimination and identification responses, according to the polarity principle.

Fig. 3. Hypothetical distributions of identification probability, topography, and latency, and of discrimination accuracy, according to the polarity principle.

Fig. 4. Observed distribution over an acoustic phonetic continuum of the identification responses /do/ and /to/. Adapted from Liberman et al.[10]

Fig. 5. Observed distributions of discrimination and identification responses for the /do/ - /to/ continuum. Adapted from Liberman et al.[10]

Fig. 6. Observed distributions of discrimination accuracy and identification probability and latency for the /do/ - /to/ continuum. Liberman et al[10] reported the first two functions; the latency function was obtained by Cross and Lane[14] from a subject who gave comparable gradients of identification probability.

Fig. 7. Observed distributions of identification topography, latency and probability for an arbitrary continuum, following conditioning with the Cross-Lane paradigm. Top: median frequency in cps (circles) and the range (vertical lines) of high- and low-pitch responses as a function of stimulus intensity for each of three subjects. The dashed horizontal lines represent the vocal pitches previously differentiated. Bottom: Response probabilities (squares) equal the ratio of the number of low pitch responses emitted to the number of stimulus presentations (10) at each intensity. The gradients for the high-pitch responses are complementary and are not shown. The hexagons represent the average latency of high- and low-pitch responses to each stimulus intensity. From Cross and Lane[14].

Fig. 8. Observed distributions of identification probability for a nonspeech control continuum, following conditioning with the Cross-Lane paradigm. The seven spectrographic patterns in the series employed by Liberman et al[10] to study /do/ - /to/ identification were inverted and then converted to sound. Each of three subjects was conditioned to respond /do/ to the stimulus with no cutback of the first formant and /to/ to the stimulus with 60 msec cutback; then each of the seven stimuli was presented 60 times in irregular order for identification. From Lane[13].

Fig. 9. Observed distributions of identification probability and discrimination accuracy for the nonspeech, inverted /do/ - /to/ continuum, following conditioning with the Cross-Lane paradigm. Each point is the relative frequency of correct discrimination in 24 presentations of pairs of stimuli separated by two or three steps in the series of seven stimuli (solid lines). This subject's identification gradient, which appears in Figure 8 (S 2), is also plotted to reveal that peaks in discrimination occur at the 'phoneme boundary' of the nonspeech stimuli. The discrimination functions predicted from the identification gradient are shown by the dotted lines. From Lane[13].

Fig. 10. Observed distribution of identification probability and discrimination accuracy before and after conditioning an aphasic adult, using the Cross-Lane paradigm, to identify the extreme stimuli in the series as /do/ and /to/ respectively. From Lane and Moore~~[16]~~.[19]

Fig. 11. Observed distributions of identification probability and latency for a nonlinguistic continuum, following conditioning with the Cross-Lane paradigm. The ninestimuli were circles, projected on a screen, with deleted sectors of 30, 34, 38 ... 62 degrees, respectively. The relative frequency of R_1 in 60 presentations of each stimulus on each of

six days is shown by the dashed lines. The gradient for R_2 is complementary and is not drawn. Frequency distributions of the latencies of both responses combined are shown at each stimulus value; the medium latency is represented by the filled triangle. From Cross, Lane and Sheppard[~~17~~ 20].

Fig. 12. Observed distributions of identification probability and discrimination accuracy for a nonlinguistic continuum, following conditioning with the Cross-Lane paradigm. Each point on the identification functions (solid lines) represents the relative frequency of R_1 in twenty-four presentations of the indicated stimulus value. The function for R_2 is complementary and is not drawn. Each point on the discrimination functions (dashed lines) represents the relative frequency of correct discriminations in thirty-six presentations of each pair of stimuli. From Cross, Lane and Sheppard[~~17~~ 20].

Fig. 13. Observed distributions of identification responses, following conditioning of a retarded child, using the Cross-Lane paradigm. Each point is the mean of 10 determinations. Adapted from Lane and Curran[~~20~~ 23].

Fig. 14. Observed distributions of identification responses, following conditioning of an acute schizophrenic adult, using the Cross-Lane paradigm. Each point is the mean of 10 determinations. From Curran[~~21~~ 24].

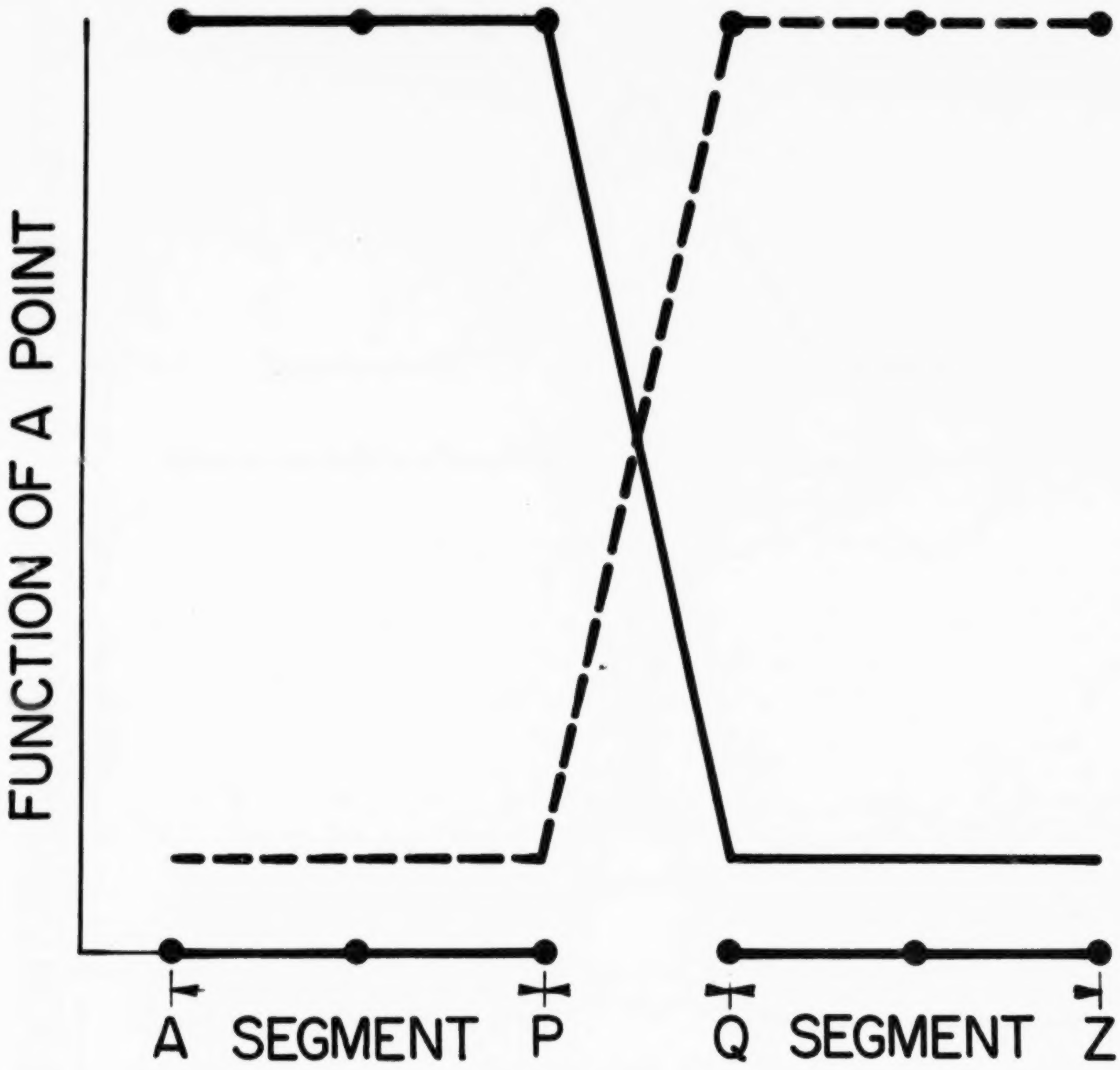

Fig. 1

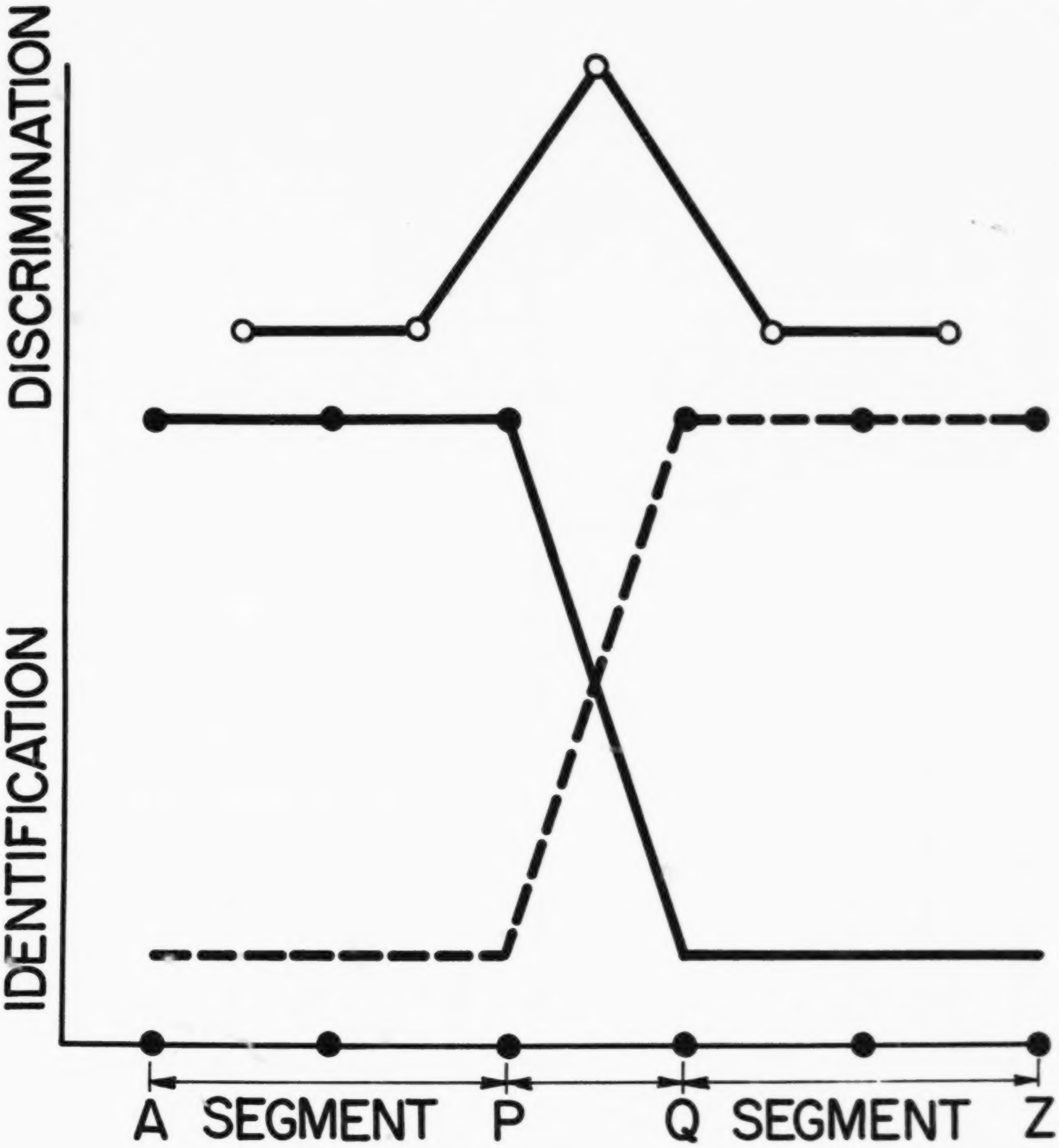

Fig. 2

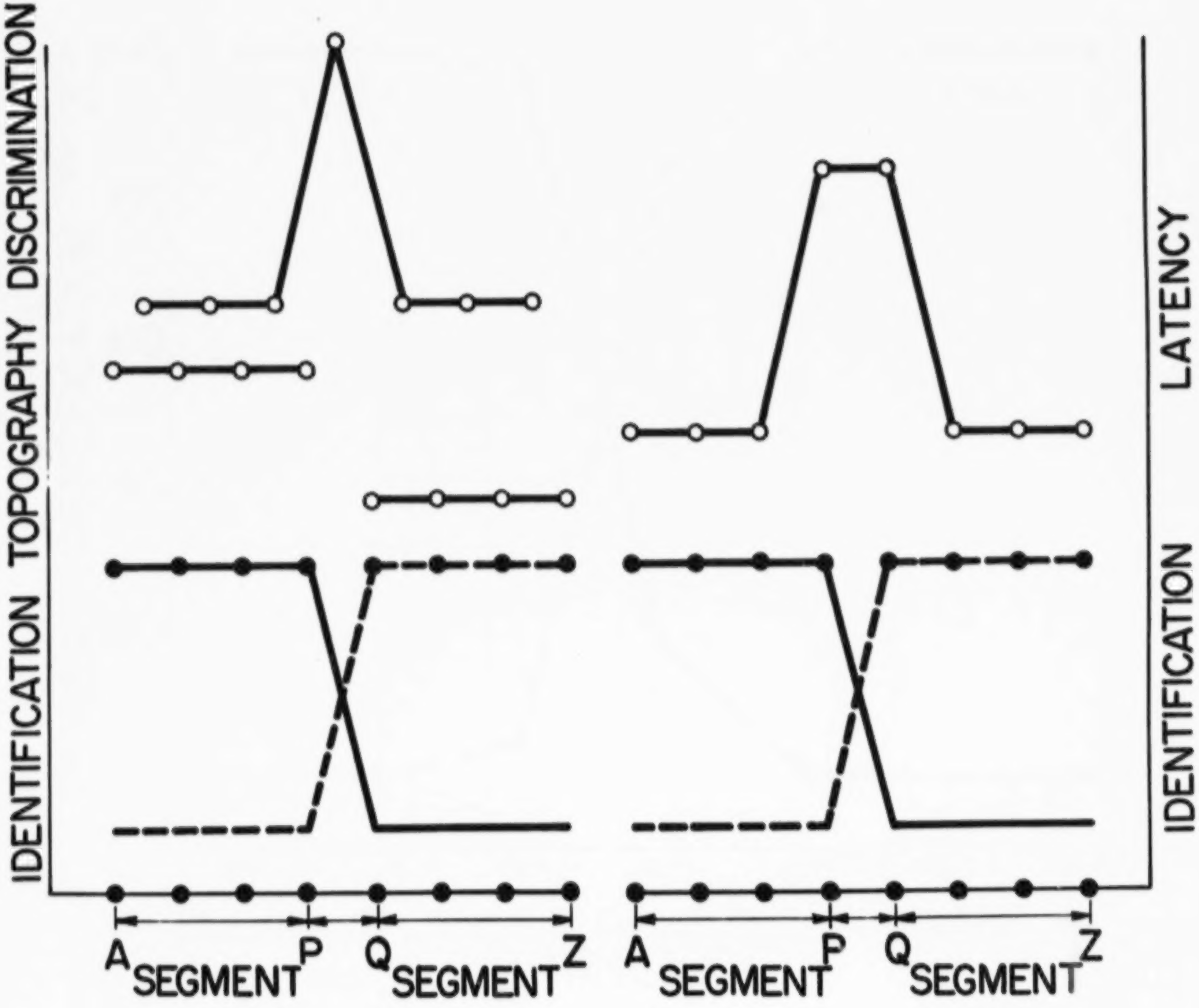
IDENTIFICATION
TOPOGRAPHY
DISCRIMINATION
LATENCY
IDENTIFICATION
A
SEGMENT
P
Q
SEGMENT
Z
A
SEGMENT
P
Q
SEGMENT
Z

Fig. 3

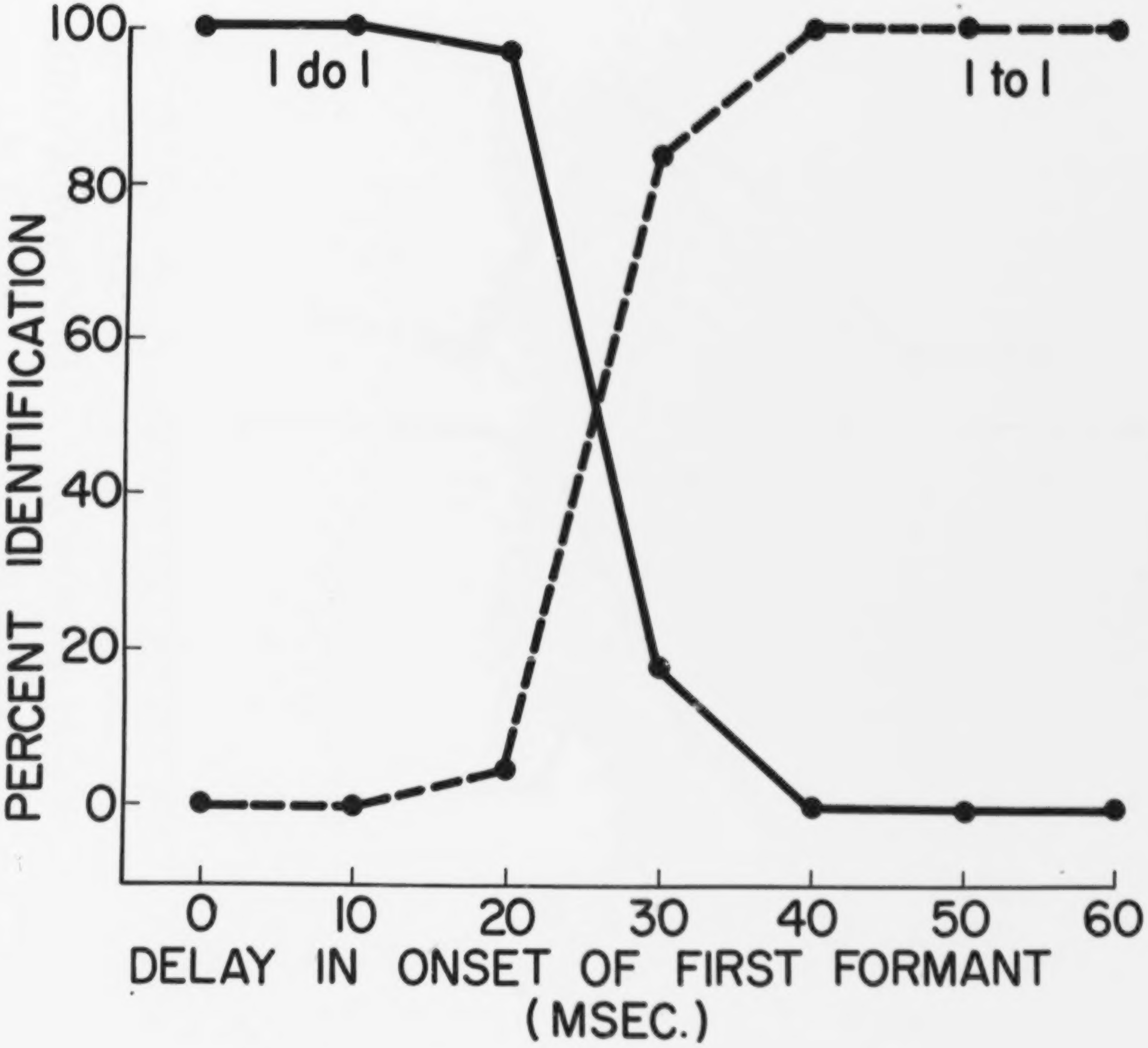
100
80
60
40
20
0
PERCENT IDENTIFICATION
/do/
/to/
0
10
20
30
40
50
60
DELAY IN ONSET OF FIRST FORMANT
(MSEC.)

Fig. 4

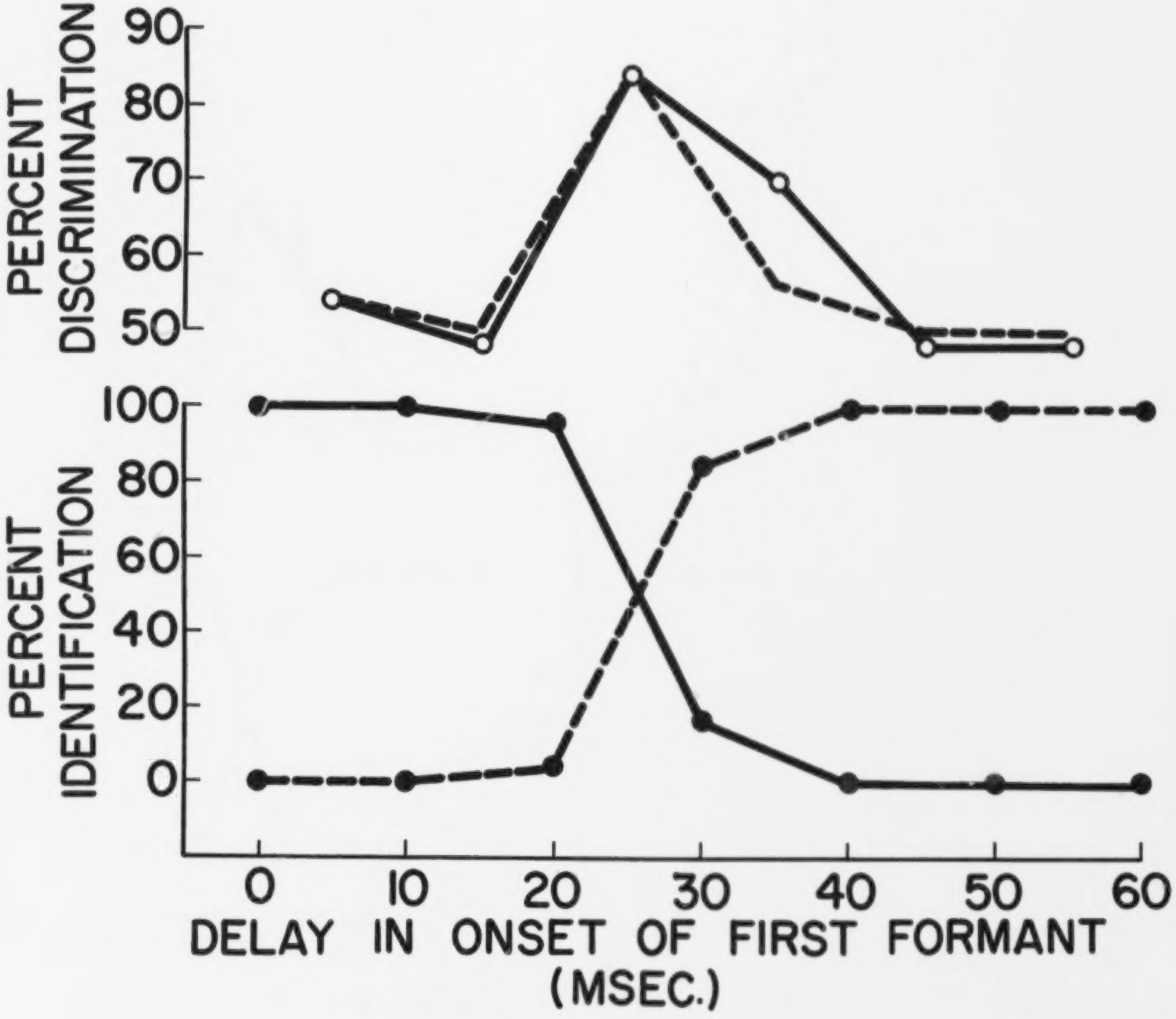
PERCENT DISCRIMINATION
90
80
70
60
50
PERCENT IDENTIFICATION
100
80
60
40
20
0
0
10
20
30
40
50
60
DELAY IN ONSET OF FIRST FORMANT
(MSEC.)

Fig. 5

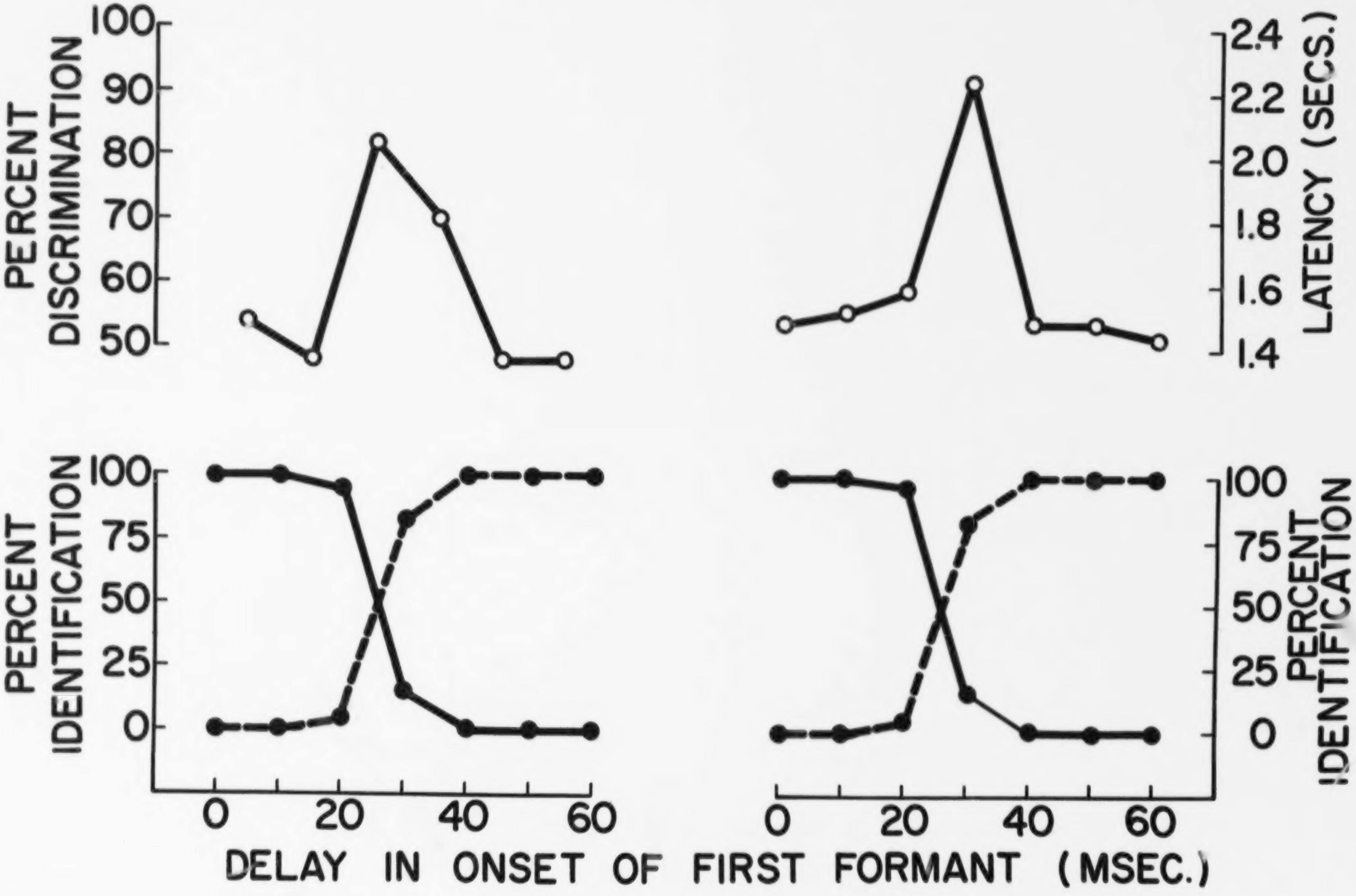
PERCENT DISCRIMINATION
100
90
80
70
60
50
LATENCY (SECS.)
2.4
2.2
2.0
1.8
1.6
1.4
PERCENT IDENTIFICATION
100
75
50
25
0
PERCENT IDENTIFICATION
100
75
50
25
0
0
20
40
60
0
20
40
60
DELAY IN ONSET OF FIRST FORMANT (MSEC.)

Fig. 6

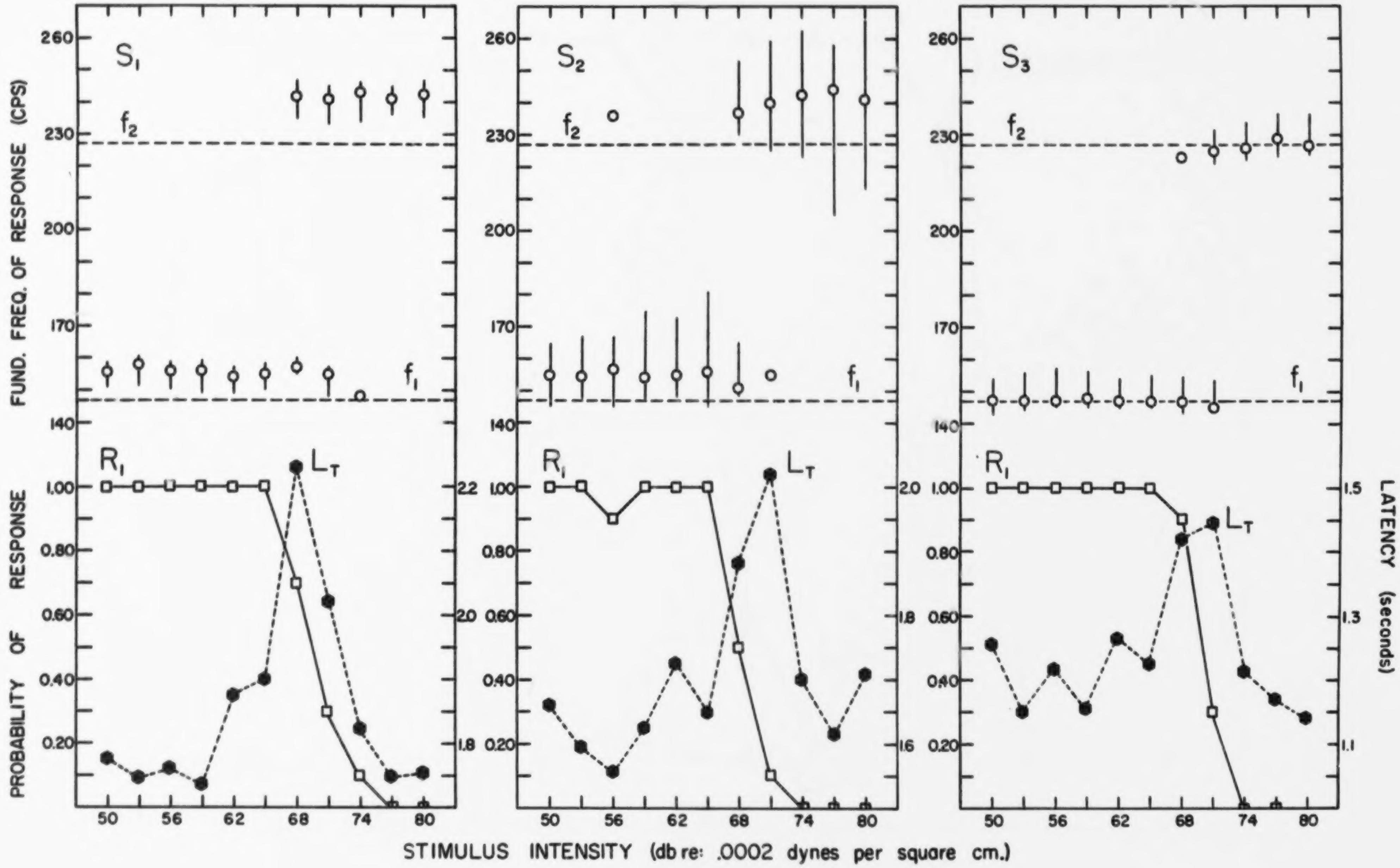
S₁
S₂
S₃
f₂
f₁
R₁
L_T
FUND. FREQ. OF RESPONSE (CPS)
260
230
200
170
140
PROBABILITY OF RESPONSE
1.00
0.80
0.60
0.40
0.20
LATENCY (seconds)
2.2
2.0
1.8
2.0
1.8
1.6
1.5
1.3
1.1
50
56
62
68
74
80
STIMULUS INTENSITY (db re: .0002 dynes per square cm.)

Fig. 7

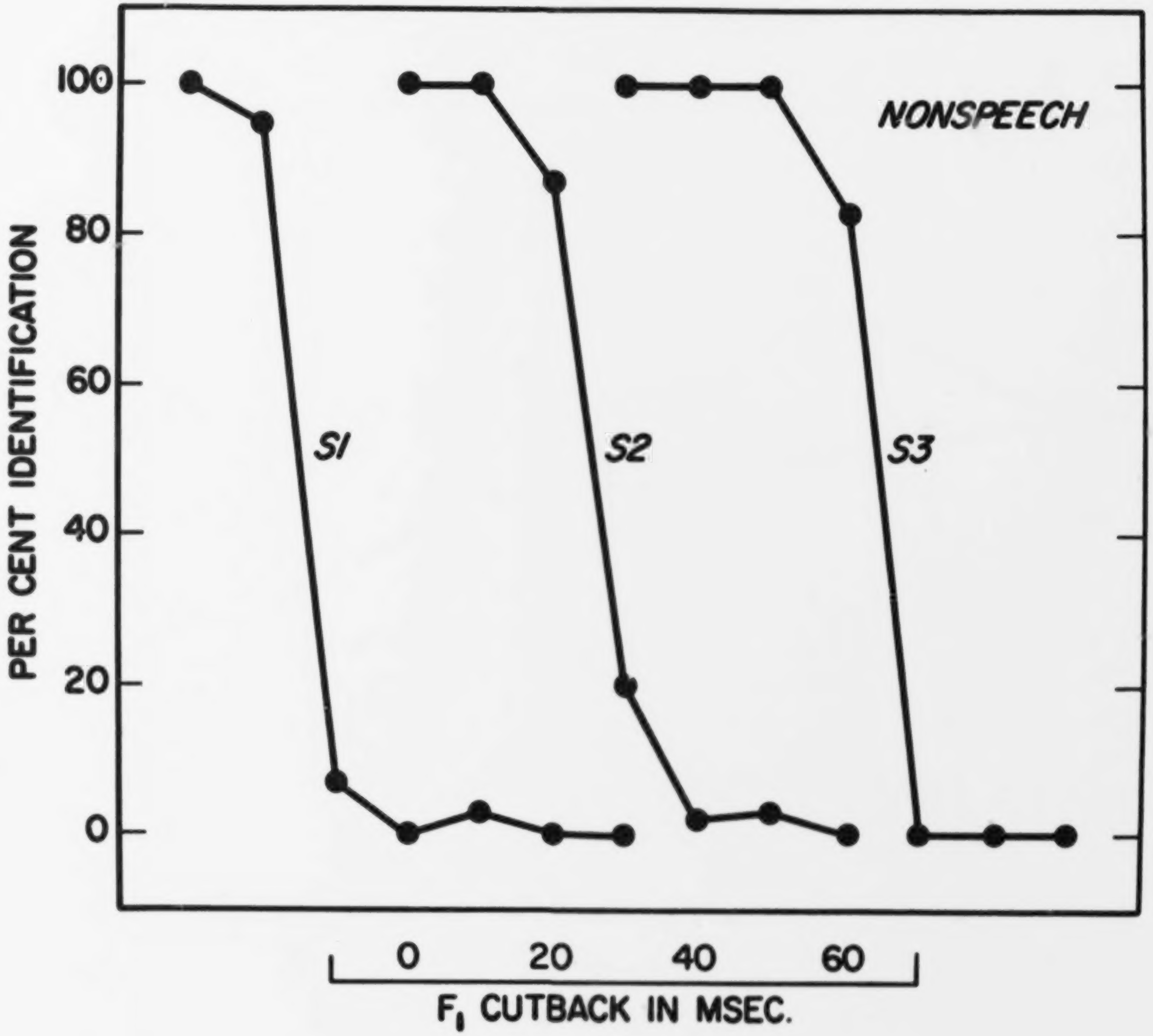
100
80
60
40
20
0
PER CENT IDENTIFICATION
NONSPEECH
S1
S2
S3
0
20
40
60
F_1 CUTBACK IN MSEC.

Fig. 8

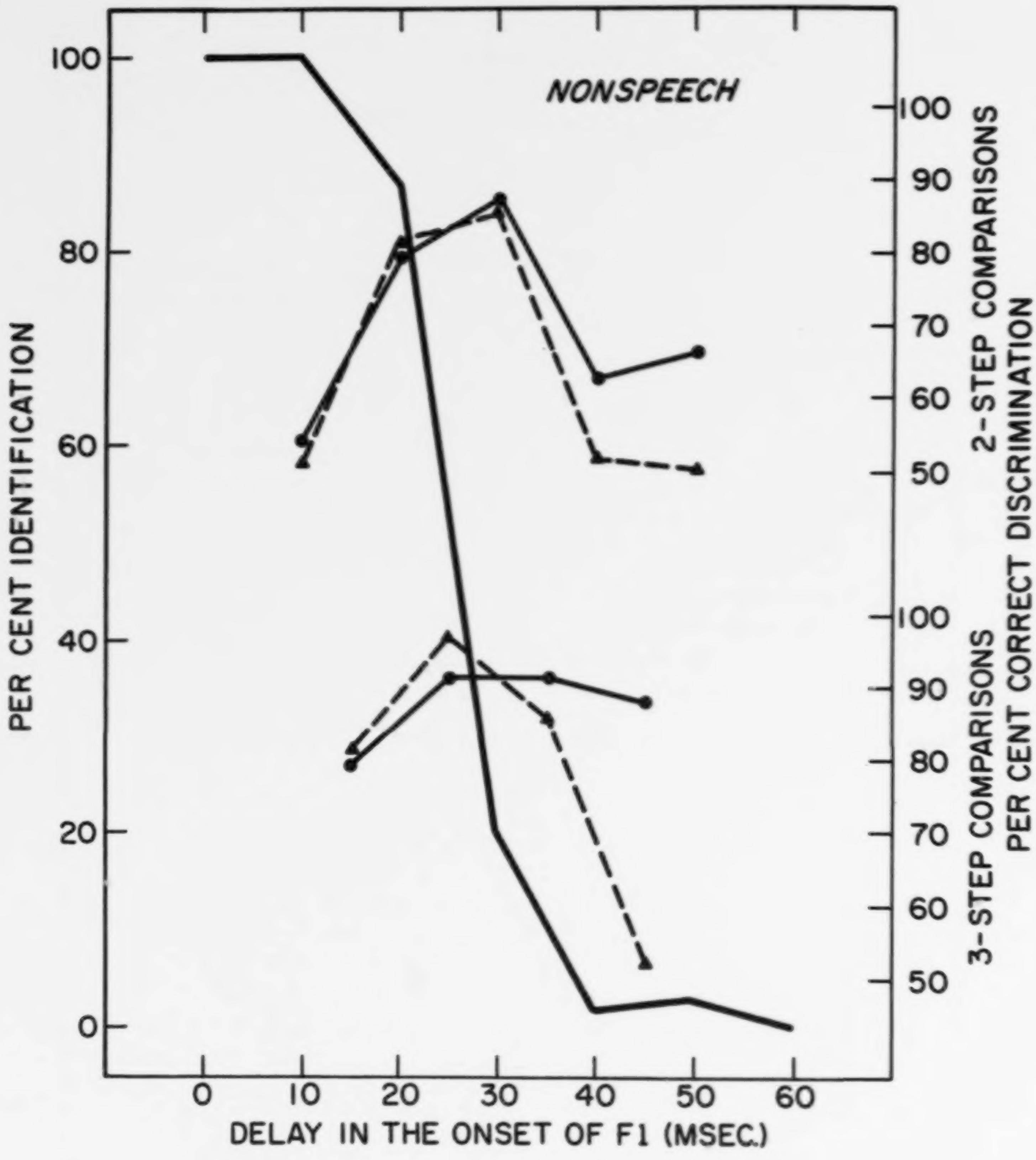
NONSPEECH
PER CENT IDENTIFICATION
0
20
40
60
80
100
0
10
20
30
40
50
60
DELAY IN THE ONSET OF F1 (MSEC.)
2-STEP COMPARISONS
3-STEP COMPARISONS
PER CENT CORRECT DISCRIMINATION
50
60
70
80
90
100
50
60
70
80
90
100

Fig. 9

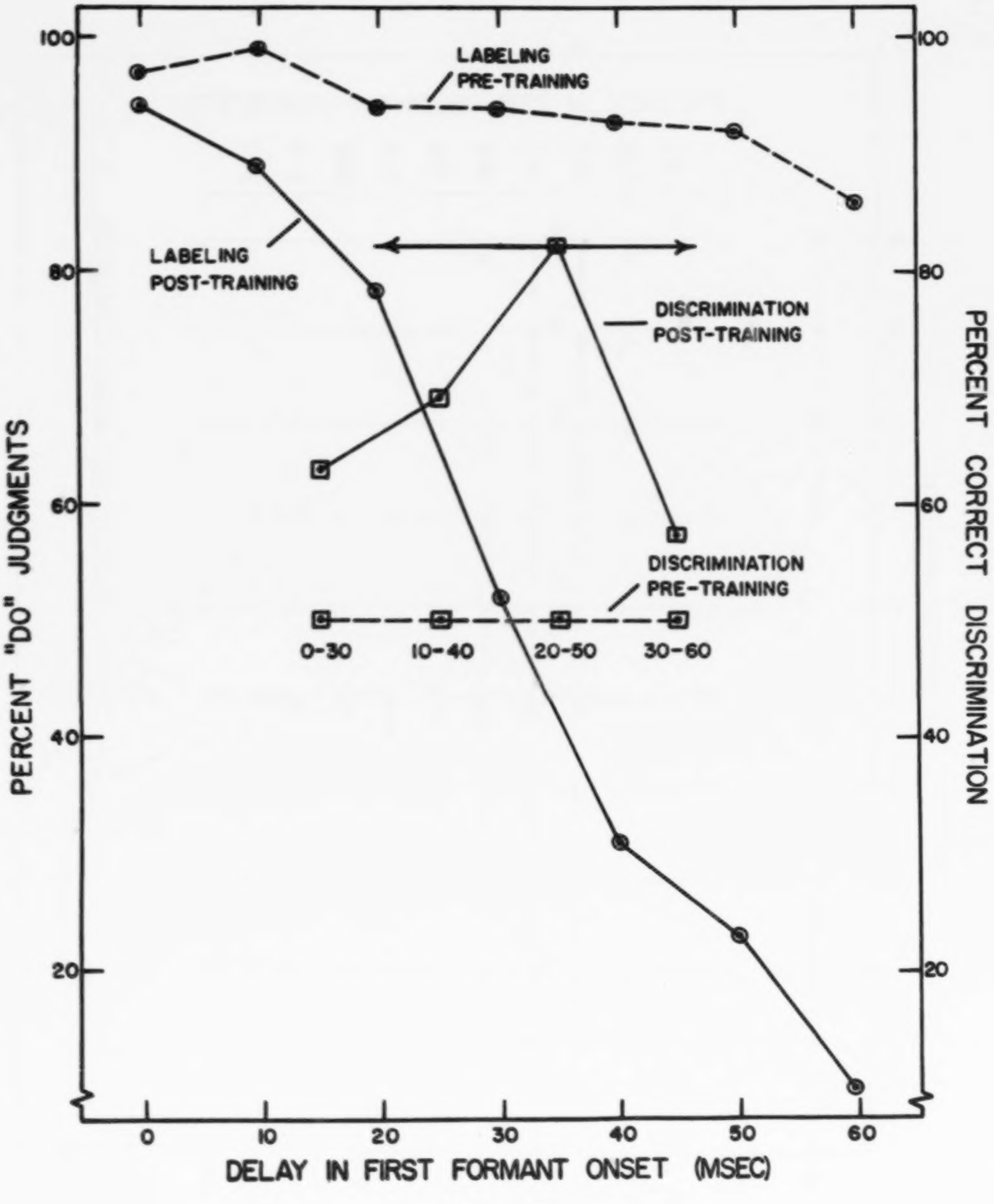
LABELING PRE-TRAINING
LABELING POST-TRAINING
DISCRIMINATION POST-TRAINING
DISCRIMINATION PRE-TRAINING
0-30
10-40
20-50
30-60
PERCENT "DO" JUDGMENTS
PERCENT CORRECT DISCRIMINATION
100
80
60
40
20
0
10
20
30
40
50
60
DELAY IN FIRST FORMANT ONSET (MSEC)

Fig. 10

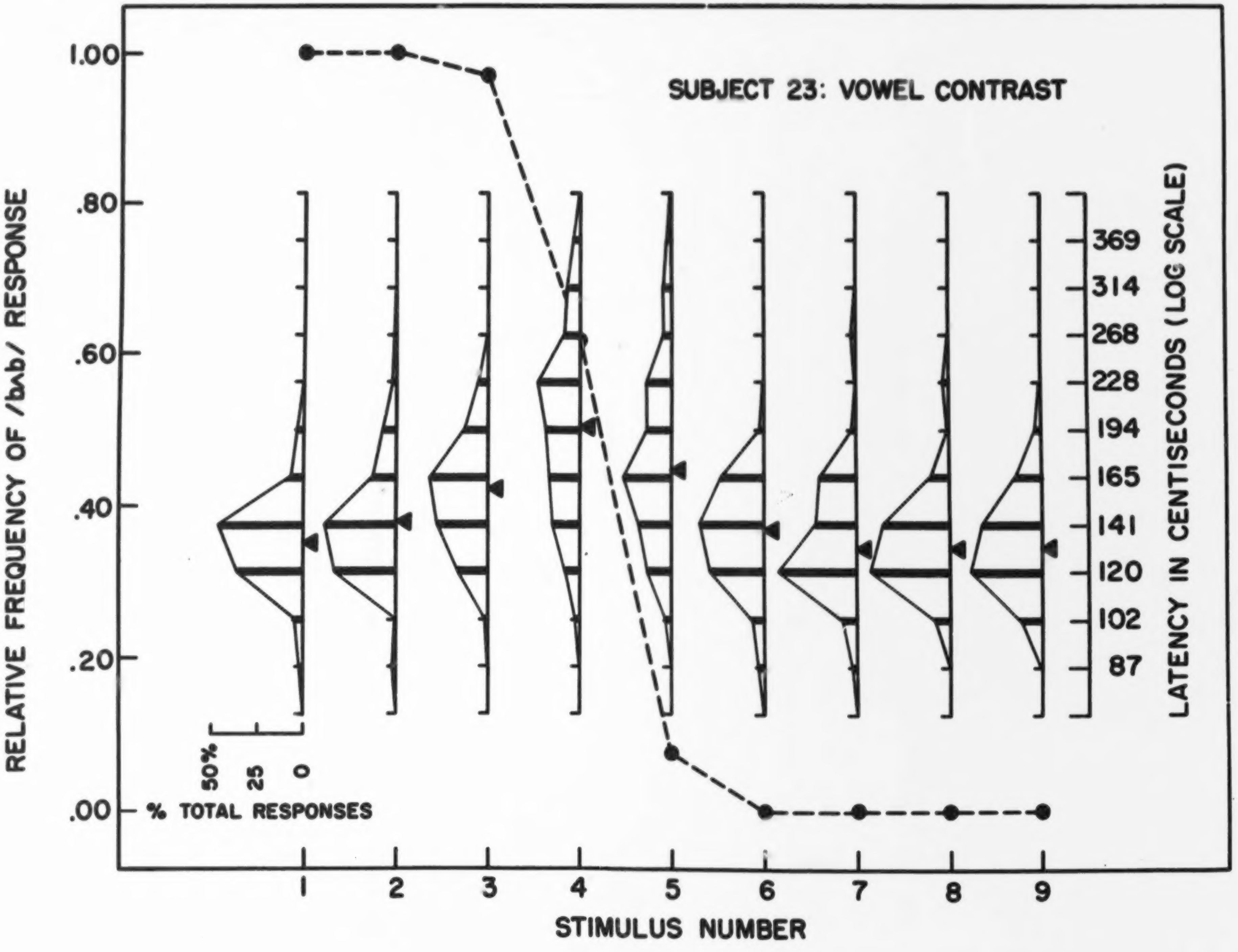
SUBJECT 23: VOWEL CONTRAST
RELATIVE FREQUENCY OF /bʌb/ RESPONSE
1.00
.80
.60
.40
.20
.00
% TOTAL RESPONSES
50%
25
0
LATENCY IN CENTISECONDS (LOG SCALE)
369
314
268
228
194
165
141
120
102
87
1
2
3
4
5
6
7
8
9
STIMULUS NUMBER

Fig. 11

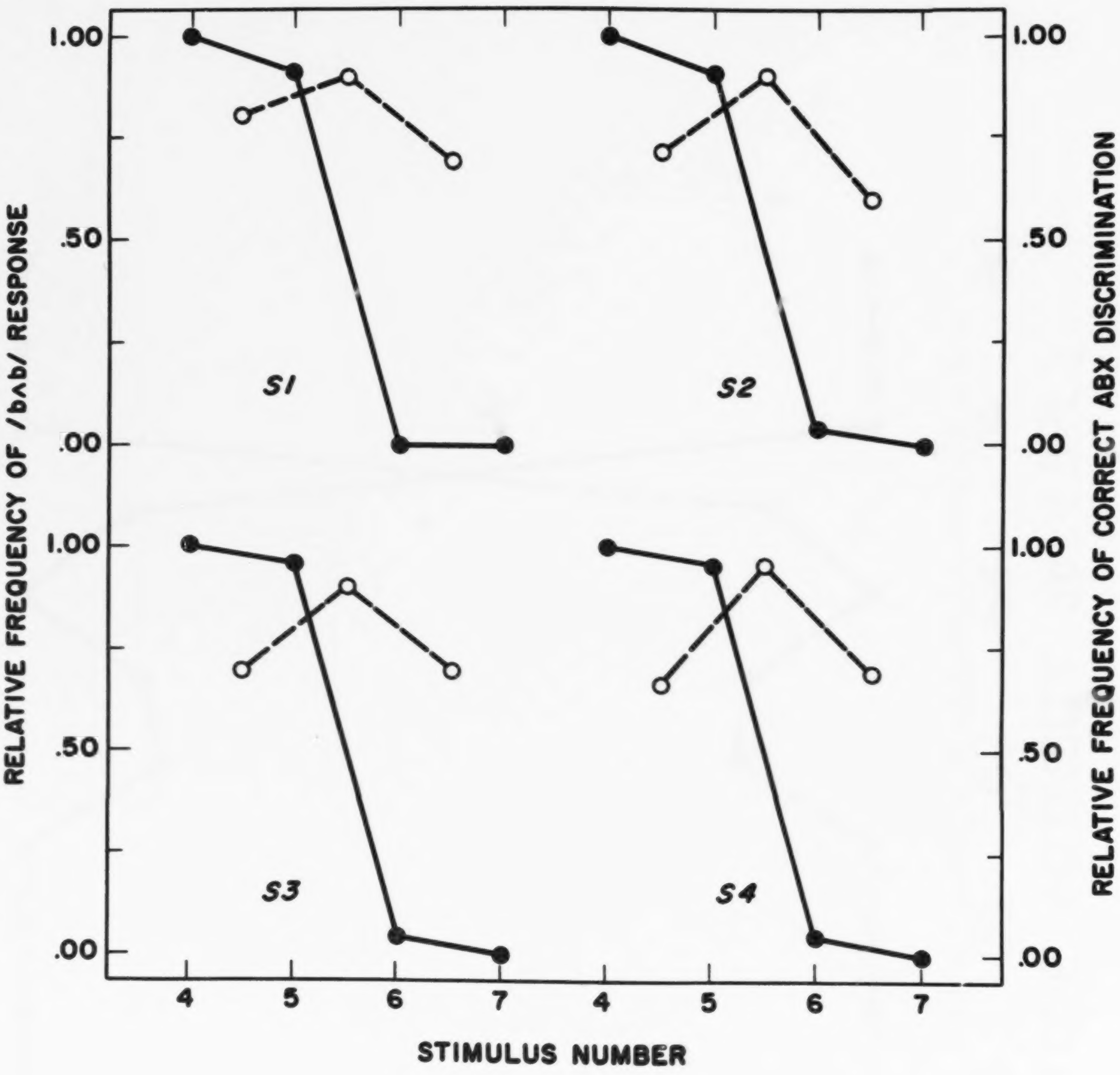
RELATIVE FREQUENCY OF /bʌb/ RESPONSE
RELATIVE FREQUENCY OF CORRECT ABX DISCRIMINATION
1.00
.50
.00
1.00
.50
.00
S1
S2
S3
S4
4
5
6
7
4
5
6
7
STIMULUS NUMBER

Fig. 12

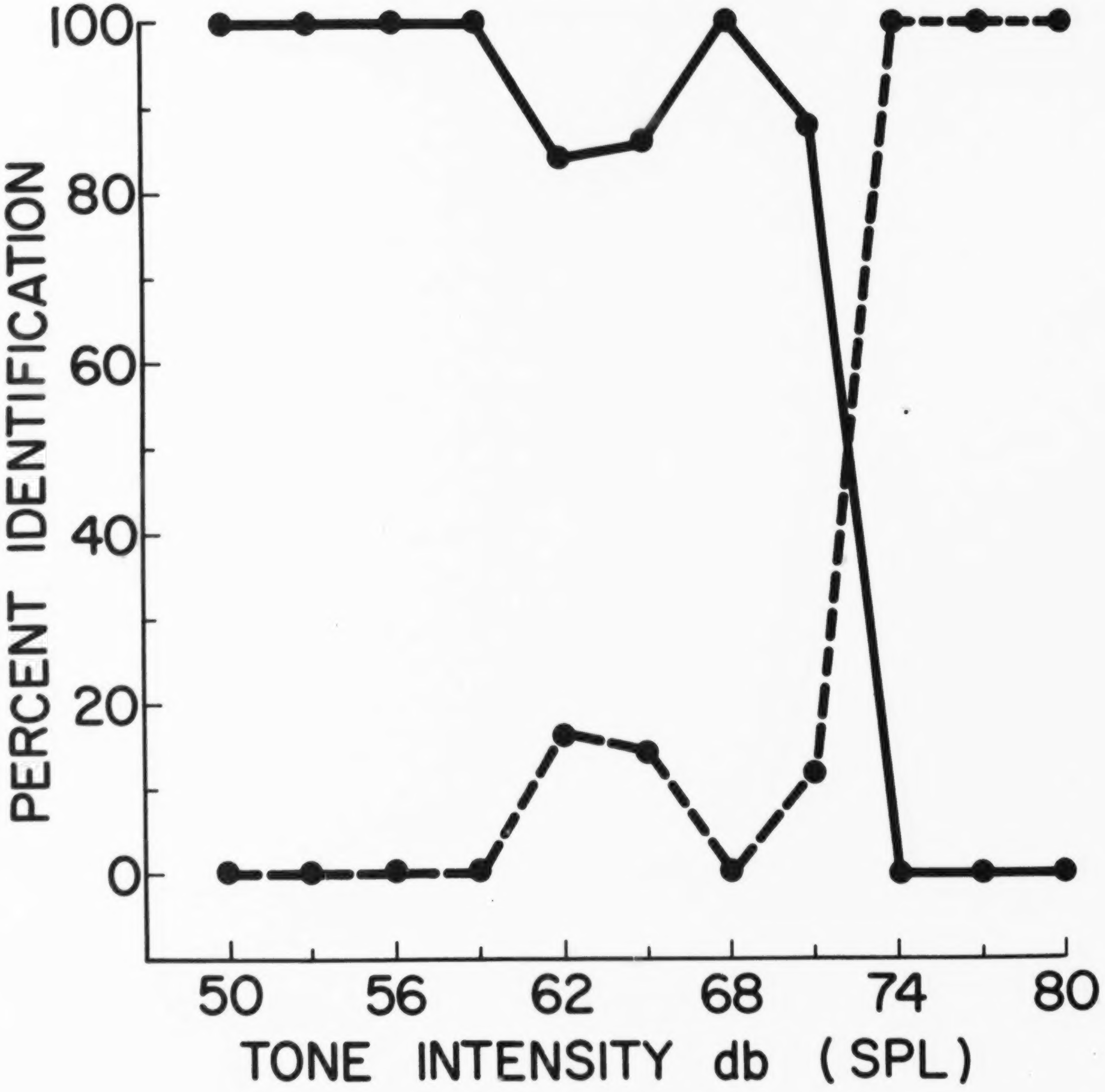
PERCENT IDENTIFICATION
100
80
60
40
20
0
50
56
62
68
74
80
TONE INTENSITY db (SPL)

Fig. 13

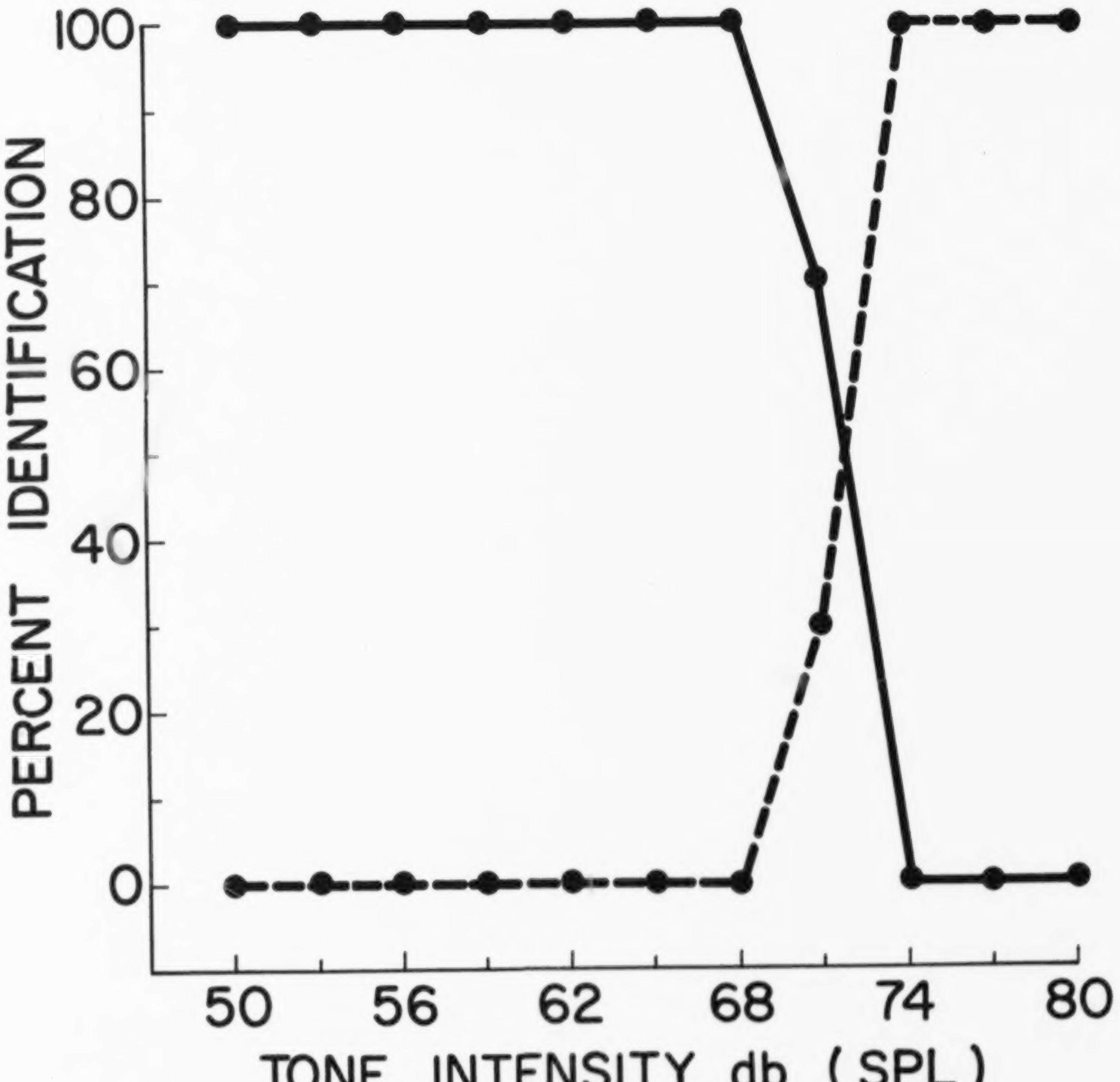
PERCENT IDENTIFICATION
100
80
60
40
20
0
50
56
62
68
74
80
TONE INTENSITY db (SPL)

Fig. 14

MODELS OF LEARNING AND METHODS OF TEACHING

Harlan L. Lane
The University of Michigan

Models of Learning and Methods of Teaching *

Harlan L. Lane
Center for Research on Language and Language Behavior
The University of Michigan

For some years I believed naively that it was mainly psychologists who had theories or models of the learning process. However, it has become increasingly clear that virtually everyone who has tried to teach another person something--and that means almost everybody--has a model of learning. Sometimes the model is implicit in the particular techniques of instruction that a person employs, but more often the model is quite explicit and anyone who asks will receive a detailed description. Most of these descriptions begin in roughly the same way: "People learn by...;" however, the descriptions all end in different ways: "People learn by practice;" "People learn by being shown;" "People learn by stops and starts;" "People learn by proceeding little by little;" "People learn by trying;" "People learn by trial and error;" "People learn by succeeding;" "People learn by imitation;" "People learn by doing it on their own;" and even "Different people learn in different ways."

Despite the apparent diversity of these principles of learning it seems that most people adhere to one of three basic models of learning. The first and most popular characterization of the learning process I will call the "sunburn" model of learning, because its fundamental premise is learning by exposure. The second model is less popular among teachers and the lay public; it is most often expressed by people engaged in personnel training in an industrial or military setting. This model derived from a short-lived amour between training supervisors and electronic engineers near the end of the war.

---

* Originally, an address read at the Indiana University-Purdue University Foreign Language Conference, March 12, 1965.

They were impressed by the analogy between methods for keeping an anti-aircraft battery on target in the course of tracking and methods for keeping a student on target in the course of learning; both, they reasoned, required continuous adjustments based on feedback of errors. This concept of the learning process may be called the "servo model," for it is fashioned after principles for controlling servo mechanisms. The third model of learning I will call the "behavioral model;" it has arisen in recent decades from research in the laboratory on the experimental analysis of human and infrahuman behavior. The behavioral model is recent and so, although it is receiving increasing attention, it is still championed primarily by experimental psychologists.

From each of these three models, the sunburn, the servo, and the behavioral, a unique set of instructional practices follows rather directly. If it is true that most of us adhere to one of these three models, it is important to consider them carefully, for the model to which you ascribe will determine your teaching technique and the materials and devices that you employ. Consider, then, the character of each of these models and their associated instructional methods. You decide which one you followed in the past and which you will choose in the future.

The basic tenet of the sunburn model is learning by exposure. The teacher, the prime source of knowledge and light (and occasionally heat), "exposes" the students to the material. The brighter students "soak it up" and eventually "see the light." Dull students, who fail to learn, are simply not sensitive or receptive.

The users of the servo model contend that learning involves more than the acquisition of knowledge by exposure; they say it requires a modification in actual behaviors. This modification is brought about, however, in a way that accords with the sunburn model: "by exposing the student to new and different material." On other counts the two models diverge again. The servo model provides for exposure contingent on the behavior of the student. Thus, there is a feedback loop from the student to the system that is presenting material to him. When feedback indicates that the communication process has failed, the presentation system may revert to material previously mastered, launch the student on a special sub-routine, or in some other way effectively sequence the material.

The behavioral model also views learning as a change in behavior. This model is based on four basic principles of behavioral control. The first is the principle of reinforcement. A reinforcer is an object or event which, if presented to a subject after he has done something, will make it more likely that he will do that sort of thing again. The principle of reinforcement provides that the likelihood and form of a response may be manipulated with reinforcement. A response is learned when it is emitted and reinforced. The second principle of the behavioral model is discrimination. This means that behavior that is reinforced most often in the presence of a particular set of stimuli will recur most often in the presence of that set of stimuli. Conversely, behavior that goes unreinforced in the presence of other stimuli will be unlikely to recur in the presence of those stimuli. The third principle, called shaping, provides that complex repertoires of behavior are acquired by learning a gradual progression of simpler behaviors. The fourth is called multiple causation. Behavior is typically the product of many controlling variables operating in concert: motivating operations, reinforcement contingencies, stimulus conditions, prior behaviors etc.

An examination of the methods of teaching that follow from these models of learning shows that the technique without equal for the sunburn theorist is the lecture method. The more students and the more exposure hours as they are called, the better. As for materials, the textbook and the reading list are helpful in exposing the student to more and different kinds of information. With regard to devices, audio-visual aids were designed for and by the sunburn theorists. Films can expose the student to more and different kinds of experiences than the lecture can alone. Television is also helpful in this regard. However, there is no provision for active responding by the students, not to mention an ongoing feedback loop to the teacher or the provision of reinforcement for correct responding.

Many language teachers operate frankly with a sunburn model of learning and their techniques are as inefficient as this characterization of learning is inappropriate. When the practices of the language classroom permit active participation by the student at all, it is often in the form of repetitive drills which are presumed to "fix" or "stamp in" the material through repetition alone. Choral work in the language classroom is usually supplemented by sonic treatments in the language laboratory; 20 minutes is generally considered the maximum safe exposure in one session. Usually the session consists of little more than simple exposure and students, like sunbathers, tend to get drowsy. In a few language laboratories there is provision for active responding by the student; and in a lesser number he may even hear his poor pronunciation played back to him repeatedly. The student is in no position to evaluate his pronunciation, however, since the novice speaker is usually a novice listener. Even if the student were able to discriminate between good and poor pronunciation, his success would have no differential effects on the immediate environment. When audible differences have no differential effects, both discrimination and production fail.

The story of language learning by exposure would have a tragic ending were it not for a deus ex machina named "language aptitude". Instead of acknowledging that the inadequacies of our students are the products of inadequate teaching methods based on an absurd model of learning, the key role in language learning is assigned to ability or endowment; this conveniently puts the process of learning beyond the teacher's control--and responsibility. Psychologists, by constructing so-called aptitude tests, have not helped to refocus attention on the critical instructional variables, those that the teacher can manipulate. Indeed, since it is invariably true that the student who takes the test is assigned an aptitude score, the presence of the score lends credence to the reality of aptitudes. This gambit has as much practical effect as lifting oneself by one's own bootstraps. Let us leave the sunburn theorist in that stance and consider the methods of teaching that follow from a servo model.

Followers of this model generally wish to beg the question of how learning takes place and are satisfied to determine whether learning has taken place from step to step in the sequence of materials presented. If there is a failure in communication the sequence of materials is altered. Therefore, the technique usually involves individual instruction, since individuals will fail at different points and for different reasons. Because individual instruction by a teacher is not feasible economically, the student is usually presented with a large quantity of textual material by a device that has the capability of adaptive sequencing. What is essentially a long lecture is interlarded with crucial questions. For mechanical convenience, these questions usually have multiple-choice answers. The student selects his answer and then turns to the indicated page in the text, or presses the corresponding button on his machine. Then he learns whether his answer is correct and, most important, which material to turn to next. Thus, the servo model, like the sunburn model, proceeds with simple exposure as a way of bringing about learning. The important difference is that followers of the sunburn model check up a few times a semester on the level of achievement of their students and then, possibly, adapt the materials for the "average" student, whereas followers of the servo model check up a few times an hour and adapt the materials accordingly for the individual student.

The method of teaching associated with the behavioral model of learning follows rather directly from the four principles of that model, enumerated earlier. The teacher, following the behavioral model, begins with the arduous task of specifying precisely the terminal behaviors desired from the student at the end of the instructional unit. Next, the student's relevant entering behaviors are specified in detail. The instructional task then becomes one of gradually building from the initial to the terminal repertoires, shaping out responses of novel form, reinforcing them in the presence of the appropriate stimulus clusters, and extinguishing them in the presence of inappropriate stimulus clusters. Each molecular bit of behavior is evoked in turn by the presentation of stimuli that will lead to the correct response, and then that response is promptly reinforced. In general, there is no provision for recurrence within a set of materials, for they are designed so that errors occur only rarely. Typically, each student is permitted to proceed at his own pace, which entails self-instruction, as in the case of the servo model.

Some of the characteristics of the method of teaching associated with the behavioral model are illustrated by an instructional system developed at the Center for Research on Language and Language Behavior for the purpose of conditioning accurate prosody in a second language. This approach to conditioning intonation, stress, and rhythm requires considerable instrumentation, but instrumentation is by no means a necessary concomitant of the instructional methods that follow from the behavioral model of learning.

In accord with the behavioral model, we sought a means for conditioning foreign language prosody that would, first, permit us to evoke increasingly complex utterances by the student and, second, provide the student with rapid, reliable and valid evaluations of his prosodic accuracy.

These pedagogical requirements led to the development of a device called SAID, the Speech Auto-Instructional Device.

SAID is an electro-mechanical instrument which performs three significant functions in conditioning prosodic accuracy in a second language. First, SAID presents to the student tape-recorded pattern sentences that are considered standards in prosodic performance. These sentences are programmed in the best known sequence for teaching prosody in the target language to a speaker of a given native language. The student is instructed to imitate the pattern sentence after he hears it. Second, SAID processes the student's imitation, and instantaneously evaluates its acceptability on the basis of its three distinct prosodic features: pitch, loudness, and tempo. Third, SAID immediately displays to the student the degree to which his imitation is unacceptable, and demonstrates how he must modify his next imitation, in the prosodic feature under consideration, in order to make the imitation more acceptable. This process of presentation-evaluation-display repeats itself until the prosody of the student's imitation is acceptable.

The first phase of the three-fold machine process is the presentation of the tape-recorded pattern sentences. The sequence of patterns presented to the student is controlled by the computer base of SAID (Fig. 1) which instructs the tape recorder to rewind when the imitation is unacceptable or to play the next pattern when the imitation is accurate.

Insert Fig. 1 about here

The second phase, evaluation, is accomplished both by the computer and by analog electronic equipment. The speech signals from the tape recorder and the student's response microphone are channeled to analog extractors of the acoustic parameters (Fig. 2).

Insert Fig. 2 about here

The fundamental frequency, the major contributor to the perception of pitch in speech is extracted by filtering each of the speech signals into two ranges within the entire fundamental-frequency spectrum of the human voice. The range containing the fundamental is selected by switching circuitry and the signal from the selected range is sent to a frequency-to-voltage converter. The output of the converter is a d.c. voltage that varies linearly with the fundamental frequency of the speech signal. In this form it is digitized and processed by the computer.

The intensity of the speech signal, the major contributor to the perception of loudness is extracted by rectifying and filtering. This produces a d.c. voltage which varies linearly with the peak instantaneous voltage of the speech signals. This d.c. voltage is digitized and processed by the computer. While processing the intensity parameter, the computer also records the time elapsed between successive intensity peaks.

The student has not been trained and therefore is not able to interpret changes in frequency, intensity, and temporal spacing accurately, since changes in these physical parameters do not produce simply corresponding changes in their apparent or psychological values. The student can, however, be expected to discriminate readily variations in the psychological effects of these physical parameters: pitch, loudness, and tempo. Threefore, the computer converts the speech parameters from linear physical units to logarithmic psychological units, according to the psychophysical functions for speech production and perception determined in earlier research.
The variation of these psychological values more accurately represents the prosodic features of speech from the point of view of the student and his

language community. While the pattern sentence is presented to the student, the values of its psychological parameters are stored in the memory of the computer. While the student is imitating the pattern, the psychological parameter values of his imitation are compared to corresponding stored values of the pattern, and three error scores are generated. These scores are compared to preset criteria of acceptability for the three features. The results of these acceptance tests determine the sequence of presentation and display.

The third phase of the machine process involves displaying the results of the comparison of pattern and imitation, and the results of the acceptance tests. The display of the comparison difference is made on a zero-center meter, so that both the direction and magnitude of error score can be monitored readily. The student's error score for one of the three features is displayed to him as his imitation proceeds (Fig. 3). A neon indicator on his panel shows him which feature is being displayed on each trial. When the student's imitation is acceptable--that is, within preset tolerance--in pitch, for example, the device advances to the next feature, loudness, and the cycle of presentation-evaluation-display is repeated for this feature. Ultimately, the imitation is acceptable with respect to all three features, whereupon the next pattern sentence in the program is presented.

Insert Fig. 3 about here

When the sunburn, servo, and behavioral models of learning are viewed from the perspective of modern psychology, the choice among them for the educator and researcher is clear. It is only the behavioral model of learning that is grounded in a general theory of the behavior of organisms, founded, in turn, on an ongoing experimental analysis of behavior. As the basic research proceeds, the model will be elaborated and modified, and so will the methods of teaching associated with it. The channel between theory and practice cannot be one-way, however. The effectiveness of our method of teaching is a measure of the adequacy of our model of learning.

Figure Captions

Fig. 1. Experimeter's console, SAID system. The photograph shows the computer base of the system with teleprinter output (left), tape-reader input, monitoring equipment, parameter extractors, and tape recorder (extreme right).

Fig. 2. Analog electronic equipment, SAID system. (Top to bottom, left to right: test oscilloscope, mixer, frequency-to-voltage converter, monitor recorder, playback amplifier, amplifiers for parameter extraction, intensity extractor, error monitor meters, computer input amplifier, voice-operated relay, power supply, switching circuitry and filter bank for extraction of fundamental frequency, switching circuitry for mode selection, tape recorder.)

Fig. 3. Student booth, SAID system. The photograph shows the loudspeakers for presentation of the pattern sentence, the microphone that receives the student's imitation, and the student console (left to right: respond light, error display, mode indicator).

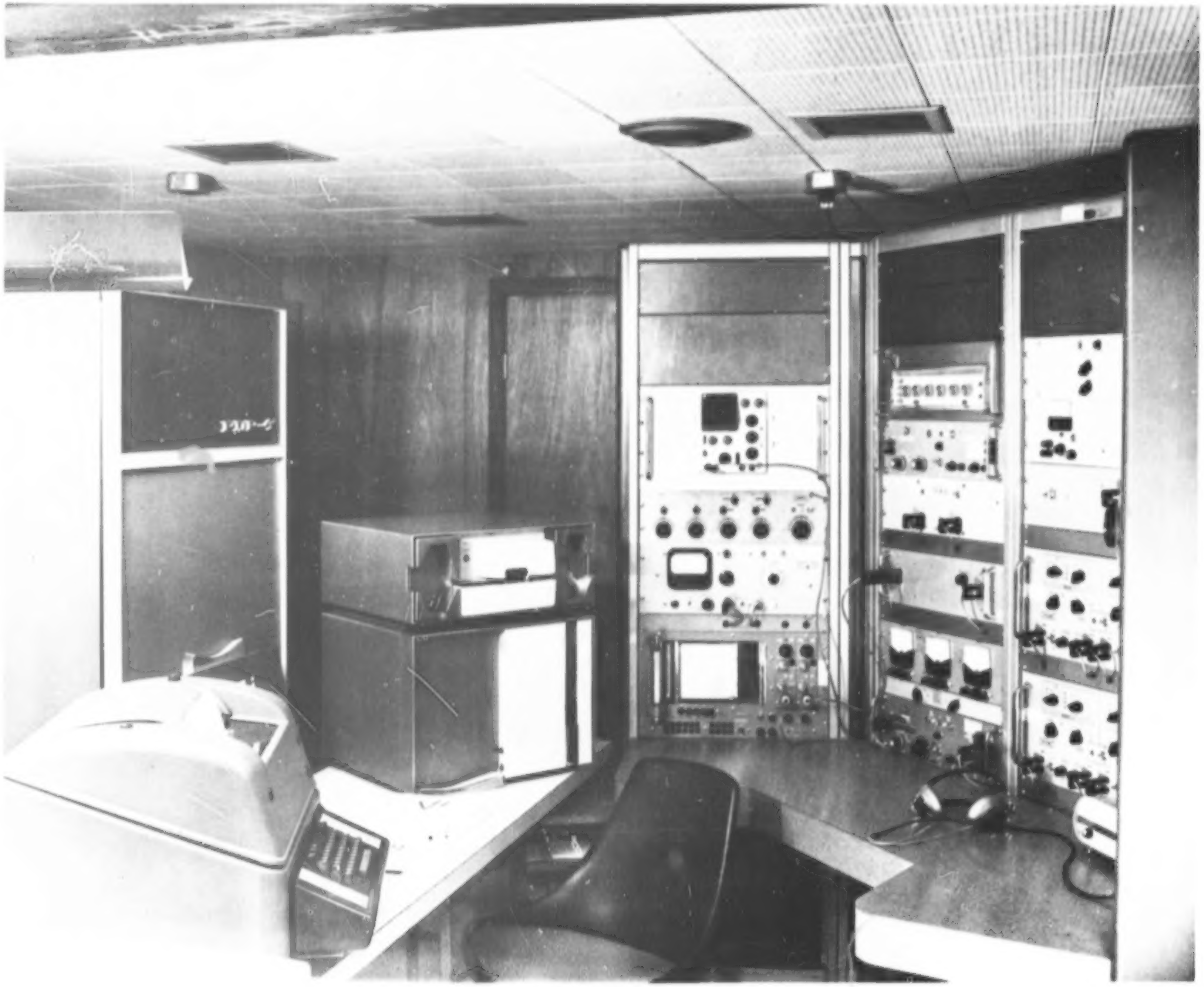

Fig. 1

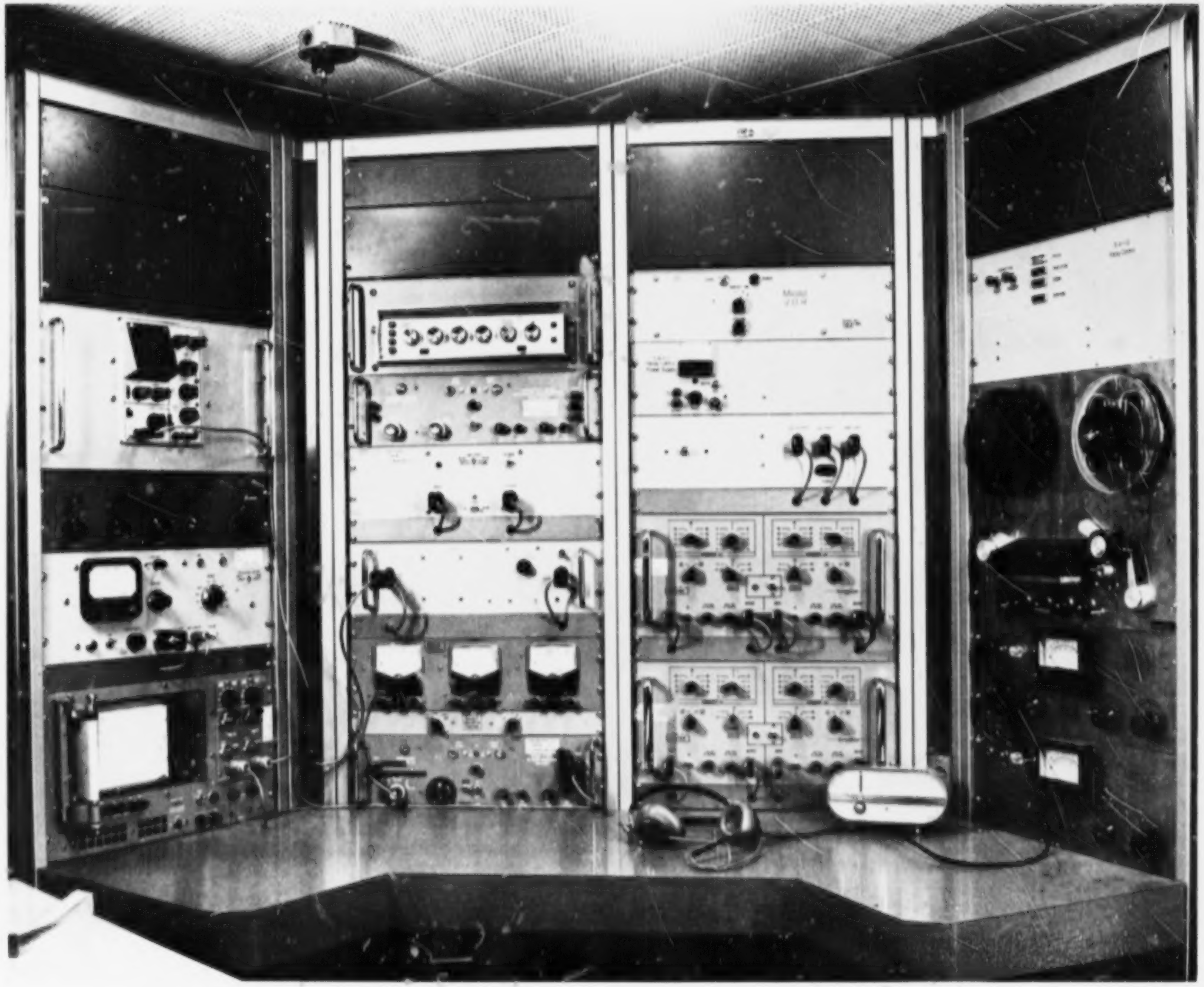

Fig. 2

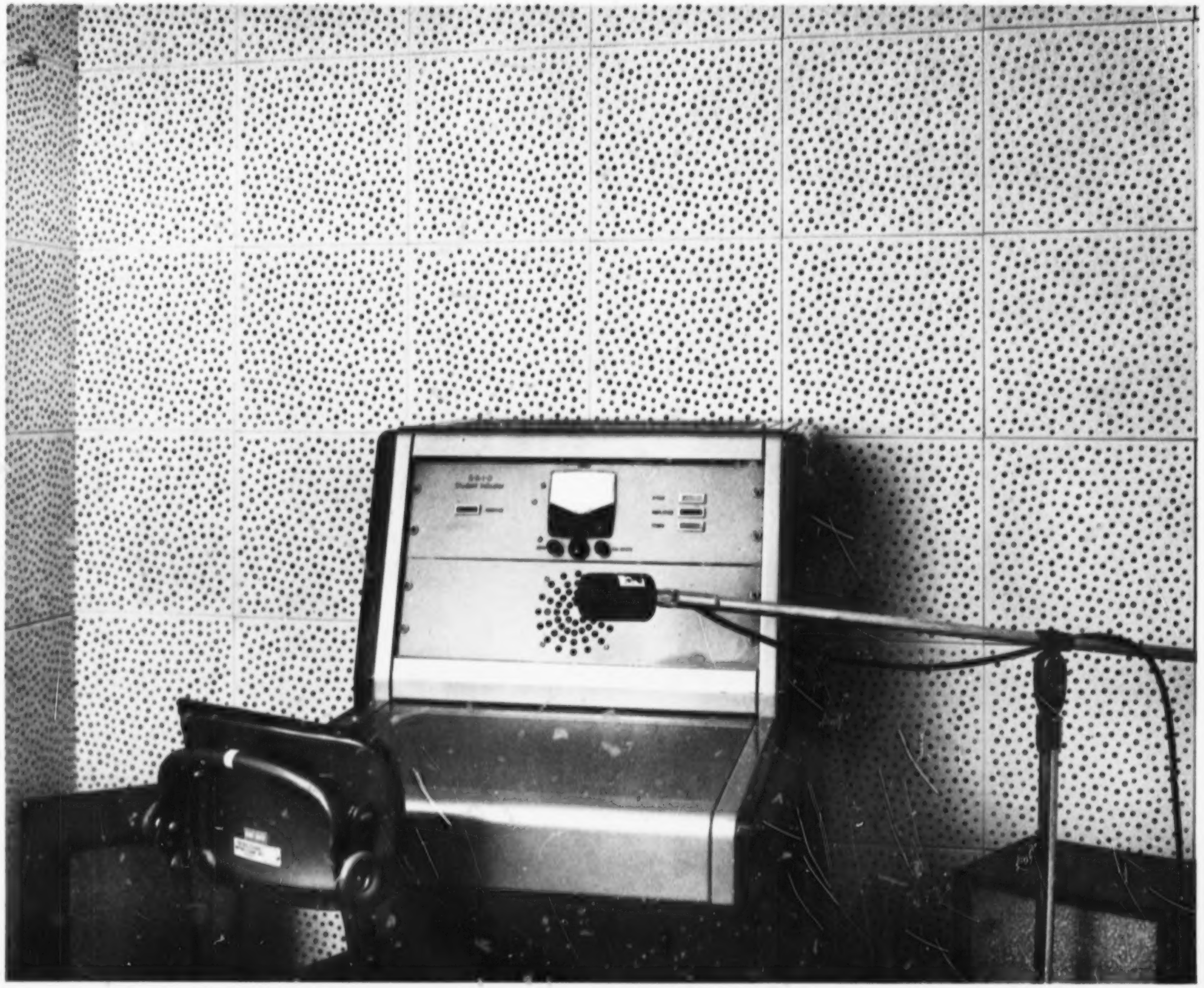

Fig. 3

RETENTION OF VOCAL TOPOGRAPHY

Harlan Lane, James Kopp and William Sheppard
The University of Michigan

RETENTION OF VOCAL TOPOGRAPHY

Harlan Lane, James Kopp and William Sheppard
The University of Michigan

Retention of Vocal Topography

Harlan Lane, James Kopp and William Sheppard
Center for Research on Language and Language Behavior
The University of Michigan

Experiments with human vocal responses have established that their temporal patterning may be controlled by schedules of reinforcement defined in units of number and time (Lane, 1960, Lane and Shinkman, 1963), and that their topography may be controlled by schedules of reinforcement defined in units of the selected acoustic parameter (e.g., duration: Lane 1964).

When selective reinforcement is applied--in the laboratory, the classroom or the clinic--to alter the vocal topography of an adult language-user, the conditioning process competes during the period of acquisition with the speaker's long history of differential reinforcement for native-language responses and it competes during maintenance with the speaker's continued successful use of native-language responses in his verbal community during the long intervals between conditioning sessions.

The two experiments to be described were designed to assess the nature and degree of the deterioration of newly-shaped vocal topography when the adult speaker was returned to his native-language community for a day.

METHOD

Apparatus

The vocal responses were transduced by a dynamic microphone, amplified, and processed by an average-speech power circuit (Peterson and McKinney, 1964). The output waveform was applied to a DC amplifier (Krohn-Hite) and then to the time-interval section of a frequency counter (Hewlett-Packard, 523D). The system was calibrated with pure-tone pulses of known duration, prepared with electronic timers and switches. The duration was read in msec and recorded by a printer and an IBM card punch.

Subjects and Procedure

Altogether 40 subjects (Ss) of both sexes were used in the two experiments. The Ss served individually in sessions lasting one to two hours.

The S was seated in an acoustically insulated chamber (Industrial Acoustics 400A) facing a panel on which a microphone, a light and a coin dispenser were mounted. These instructions were read to S:

> "In this experiment you can earn money simply by saying [u] whenever the light in front of you flashes. You will know that you have earned money when a penny falls into the tray in front of you. At the end of the experiment, you may take home all of the pennies you have earned. If you attend all sessions you also receive a bonus. Occasionally, the light will remain off for a period of time. When the light is off, please do not respond. Please make no other sounds whatever besides [u] throughout the experiment."

An electronic timer flashed the stimulus light at 5-second intervals. The experimenter entered the duration of each response on a desk calculator; this allowed him to compute, as the session proceeded, the mean and average deviation of response duration for a given number of preceding responses. The experimenter presented reinforcement manually.

There were two types of reinforcement schedules in effect at various times as the experiment progressed. During continuous reinforcement (CRF), all responses were reinforced, regardless of their duration. During selective reinforcement (shaping), only those responses whose duration exceeded a "phase criterion" were reinforced. The criterion for each phase equalled the mean duration of the ten consecutive reinforced responses that terminated the preceding phase, plus their average deviation.

The CRF phases yielded both a (1) baseline measures for determining response duration before differential reinforcement, and a criterion duration for use in the initial shaping phase, and (2) probe measures for determining response duration after shaping.

At the end of the experiment, the S was given a three-item questionnaire. The response to each item was completed before the next was presented. The questions were: 1) What do you think was the point of the experiment? 2) Did you know what it was about your responding that earned pennies for you? 3) Did you know that you received pennies depending on the length or duration of your response?

EXPERIMENT I

The purposes of this experiment were:

(1) To assess the test-retest reliability of the CRF phase as a measure of the average response duration of individual Ss prior to differential reinforcement (Groups Ia, Ib);

(2) To determine the extent to which durations measured during a CRF phase, immediately following shaping, are comparable to durations measured during the terminal portion of the shaping phase--that is, to provide a baseline measure of the change in the repertoire caused by probing under CRF (Group Ia);

(3) To assess the amount of retention for responses of a given duration, shaped by differential reinforcement, after the subject has returned to the language community for approximately 24 hours. (Group Ib).

The sequencing of CRF and shaping phases for both groups is shown in Table 1. Ten Ss were assigned randomly to each of two groups.

The treatments of the groups differed only in the delay between the last shaping phase and the last CRF phase: there was no delay for Group Ia and a 24-hour delay for Group Ib.

The first two CRF phases constituted replicative "pre-test" conditions to assess the day-to-day variability in the Ss vocal durations, given only normal experience in the language community. The third CRF phase yielded data for the computation of the criterion duration (mean plus average deviation of duration) for reinforcement in the first shaping phase.

Phases IV - VII comprised four shaping phases. Each phase terminated when ten consecutive reinforced responses were emitted; a new criterion was computed, as described above, and a new phase began without interruption.

The last CRF phase constituted a "post-test" condition to assess the extent to which the topography had retrogressed after the S had been returned to the language community for approximately 24 hours.

RESULTS AND DISCUSSION

The changes in mean response duration during replication of the pre-test and in shaping were quite comparable for the two groups (Table 2). The average durations in CRF-1 and CRF-2 (20 responses each) are virtually identical (±6 msec) within groups, despite the fact that the test-retest interval was approximately 24 hours. Over the four shaping phases, the mean duration of the last 10 responses per phase increased to well over 600 msec for both groups; relative variability remained nearly constant, indicating the differential reinforcement of longer duration led to an upward shift in the total response distribution. The average durations during the CRF phases agree well with the comparable measurements reported by Lane (1964) and with the value reported by Ilse Lehiste and Gordon Peterson (1960) for the average duration of [u] as a syllable nucleus in American English. The coefficient of variation, ranging around 11 percent (± 3%), agrees exactly with prior findings and with an initial estimate of the Weber fraction for produced vocal duration (Lane, 1964).

When CRF conditions were instituted immediately after the fourth shaping phase, the average vocal duration decreased 13 percent. If the CRF post-test was delayed 24 hours, however, the retrogression toward the baseline was 38 percent. It is instructive to consider this retrogression relative to the change in vocal duration initially effected by shaping--that is, to ask what fraction of the ground gained by shaping had been lost. Specifically, a "retention index" was computed as follows:

$$\text{Retention in percent} = 1 - \frac{\text{Shape N - post-test}}{\text{Shape N - pre-test}} \quad \text{X} \quad 100$$

where post-test = mean duration during CRF 4

pre-test = mean duration during CRF 1

Shape N = mean duration of the last 10 responses of the last complete shaping phase.

In the absence of the interpolated return to the language community, Group Ia showed an average relative decrement of 25 percent or retention of 75 percent. The retention of the typical S was greater at the outset of the post-test and then quickly leveled off near the average value. This "drift" may be attributed to retrogression toward the first-language baseline which occurs as soon as differential reinforcement is discontinued for a few responses, or it may be attributed to the operation of a "law of least effort"; the results of experiment II, described below, clearly favor the former interpretation.

The Ss in Group Ib sustained an average relative loss of 52 percent (retention index = 48 percent). From one perspective, this degree of retention can actually be considered substantial, since there were 24 hours of intervening return to the language community (presumably involving differential reinforcement for short-duration responses), and as much as 25 percent loss is to be expected even when there is no interpolated delay (Group Ia). Individual data corroborate the trends shown by the averaged group data. The durations of the responses emitted during the experiment by each of two representative Ss are shown in Fig. 1 and Fig. 2. Neither of these Ss described the contingencies of reinforcement on their questionnaire at the completion of the experiment, although their response durations increased over 300 percent. Figure 3 shows the average duration of responding by each S during the pre-test, the termination of the last shaping phase, and the post-test, as well as the corresponding value of the retention index.

The results of the questionnaires reveal that about 50 percent of the Ss could correctly describe the contingencies of reinforcement at the end of the experiment. Of the Ss who could not, about half discriminated that topography of the vocal response was involved (pitch, vocality, etc.). A point-biserial correlation coefficient of + .66 ($P < .01$) was obtained between positive versus negative questionnaire scores and mean response duration during the last shaping phase.

EXPERIMENT II

This experiment essentially replicates the treatment of group Ib, with the addition of a final retraining condition, designed to assess the extent to which there is savings in reinforcement control of vocal responding following the return to the language community.

Twenty Ss, assigned randomly to two groups of 10 each, served individually in two sessions spaced 24 hours apart. The experimental protocols for the two groups are summarized in Table 3. The procedures during the CRF and shaping phases were the same as in Experiment I. On the first day, E administered on CRF phase and two complete shaping phases. On the second day, E administered one CRF phase and as many shaping phases as it was possible to administer until the S had emitted the same number of responses during shaping on day 2 as he had on day 1.

For one of the groups (II a), successive criteria during shaping were gradually increased by adding the average deviation to the mean duration of the terminal responses of the preceding phase (shape up). For the other group (II b), the criteria were gradually decreased by subtracting the average deviation from each terminal mean duration (shape down).

RESULTS AND DISCUSSION

In many respects Experiment II replicates the procedures and the findings of Experiment I. The baseline durations, the measures of variability, the change in duration effected by shaping and the value of the retention index after interpolated delay are comparable for all groups. It is particularly interesting to note that Ss whose vocal durations are shaped to longer values show a decrease of 31 percent toward the language baseline after the return to the community, or a comparable retention of 39 percent.

Once again, the statements based on averaged group data reflect individual performance well. Figure 4 plots the durations of responses emitted by a representative *S* in group IIa (shape up) and Fig. 5 presents these findings for a typical *S* in group IIb (shape down). Neither of these *Ss* was able to describe on the questionnaire the contingencies of reinforcement during the experiment. Figures 6 and 7 present the average duration of response by each *S* in groups IIa and IIb, respectively, during the pre-test, the termination of the second shaping phase, the post-test, and the termination of the final shaping phase, as well as the corresponding measures of retention and savings.

The savings enjoyed in shaping on day 2, as a result of the initial shaping on day 1, could be assessed by comparing the ground covered by differential reinforcement on days 1 and 2, provided that the number of responses on the two days was kept equal. In particular, the extent of savings was taken to be the percent of the total change in duration on day 2 remaining after the absolute change in duration on day 1 was subtracted from the absolute change in duration on day 2.

$$\text{Savings in percent} = \frac{(\Delta D,\ \text{day } 2) - (\Delta D,\ \text{day } 1)}{\Delta D,\ \text{day } 2} \times 100 =$$

$$= \frac{(\text{Shape } N' - \text{post-test}) - (\text{shape } 2 - \text{pre-test})}{(\text{Shape } N' - \text{post-test})} \times 100 =$$

$$= \frac{1 - \text{Shape } 2 - \text{pre-test}}{\text{Shape } N' - \text{post-test}} \times 100$$

where:

Pre-test = mean duration during CRF on day 1

Post-test = mean duration during CRF on day 2

Shape 2 = mean duration of the last 10 responses of the second (last shaping phase on day 1

Shape N' = mean duration of the last 10 responses of the last complete shaping phase on day 2.

Group IIa had an average savings of 28 percent during shaping on day 2 (Table 4). For Group IIb the savings index was -48 percent; this means that, during differential reinforcement of shorter response durations, there was less of a decrease from CRF to the last shaping phase on day 2 than on day 1. This outcome reflects the fact that response durations were probably approaching a lower bound during "shape down." The presence of this lower bound is also indicated by the high inverse correlation (rho = -.63, $p < .05$) obtained between amount of retention and amount of savings for group IIb contrasted with the zero correlation (rho = .02, N.S.) between retention and savings for Group IIa. That is, in the case of shaping down, poor retention meant retrogression away from the area of minimal response duration and hence more room for reconditioning on day 2, yielding, therefore, higher measures of savings. Still further evidence for the operation of a lower limit on vocal duration can be seen in the fact that point-biserial correlations for questionnaire scores and mean response duration during the terminal 10 responses of the last shaping phase were +.75 ($p < .01$) and +.09 (N.S.) for Groups IIa and IIb, respectively. Although the ability to discriminate the contingencies of reinforcement normally allowed greater efficacy during training, it may be deduced that the existence of a physiological limit cancelled this advantage for Group IIb.

Considering the groups as a whole, seven of the questionnaires submitted by Ss in group IIa revealed that they had verbally discriminated the contingencies of reinforcement, whereas only four of the Ss in group IIb demonstrated such a discrimination. As in Experiment I, about half of the other Ss discriminated that a vocal paramenter was involved.

Table 1. Sequences of CRF and Shaping Phases in Experiment I (Retention)

PHASE	I	II	III	IV	V	VI	VII	VIII
Schedule of Reinforcement	CRF-1	CRF-2	CRF-3	Differential Reinforcement	Differential Reinforcement	Differential Reinforcement	Differential Reinforcement	CRF-4
Purpose	Pretest 1	Pretest 2	Determine First Criterion	Shaping Longer Durations	Shaping Longer Durations	Shaping Longer Durations	Shaping Longer Durations	Post-Test
Criterion for Termination	20 S^R	20 S^R	10 S^R	10 Consec. S^R	10 Consec. S^R	10 Consec. S^R	10 Consec. S^R	20 S^R
Day Group IIa (Immediate Post-Test)	1	2	2	2	2	2	2	2
Day Group Ib (Post-Test Delayed 24 Hrs.)	1	2	2	2	2	2	2	3

Table 2

Effects of CRF, Shaping and Return to the Language Community

on the Average Duration of Vocal Responses (Experiment I)

PHASE	I	II	VII	VIII
Schedule of Reinforcement	CRF-1	CRF-2	Differential Reinforcement	CRF-4
Purpose	Pretest 1	Pretest 2	Shaping Longer Durations	Post-test
		Group Ia (N=10)		
Day	1	2	2	2
Mean Duration	245	239	624	540
Percent Change		-2	+161	-13
Coeff. of Variation	.14	.10	.11	.11
Retention Index (%)			75	
		Group Ib (N=10)		
Day	1	2	2	3
Mean Duration	261	266	670	416
Percent Change		+2	+152	-38
Coeff. of Variation	.09	.09	.08	.11
Retention Index (%)			48	

Table 3. Sequences of CRF and Shaping Phases in Experiment II (Retention and Savings)

Group IIa-Shaping Longer Durations

Group IIb-Shaping Shorter Durations

PHASE	I	II	III	IV	V	N
Day	1	1	1	2	2	2
Schedule of Reinforcement	CRF-1	Diff S^R of Responses Exceeding Criterion	Diff S^R of Responses Exceeding Criterion	CRF-2	Diff S^R Responses Exceeding Criterion	Diff S^R of Responses Exceeding Criterion
Purpose	Pre-test	Shaping longer or shorter durations	Shaping longer or shorter durations	Post-test	Shaping longer or shorter durations	Shaping longer or shorter durations
Criterion for Termination	20 S^R	10 Consec. S^R	10 Consec. S^R	20 S^R	10 Consec. S^R	Same No. of Responses as Were Emitted on Day 1

Table 4. Effects of Shaping, Return to the Language Community, and Reconditioning on the Average Duration of Vocal Responses (Experiment II)

PHASE	I	II	IV	N
Schedule of Reinforcement	CRF-1	Differential Reinforcement	CRF-2	Differential Reinforcement
Purpose	Pre-test	Shaping	Post-test	Shaping
Day	1	1	2	2
		Group IIa (N=10 shape up)		
Mean Duration	250	563	388	1634
Percent Change		+114	-31	+337
Coeff. of Variation	.12	.08	.13	.08
Retention Index (%)		44		
Savings (%)			28	
		Group IIb (N=10 shape down)		
Mean Duration	239	167	203	156
Percent Change		+30	-22	+22
Coeff. of Variation	.11	.06	.09	.06
Retention Index (%)		39		
Savings (%)			-48	

FOOTNOTES

1. This study of retention of vocal topography developed from a suggestion by Dr. D. Lee Hamilton. Thomas Anderson assisted in the conduct of the research.

REFERENCES

Lane, H. L. 'Differential reinforcement of vocal duration.' J. exp. anal. Behav., 1964, 7, 107-115.

Lane, H. L. 'Temporal and intensive properties of human vocal responding under a schedule of reinforcement.' J. exp. anal. Behav., 1960, 3, 183-192.

Lane, H. L. and Shinkman P. G. 'Methods and findings in an analysis of a vocal operant.' J. exp. anal. Behav., 1963, 6, 179-188.

Lehiste, Ilse and Peterson, G. E. 'Some acoustic properties...'

Peterson, G. E. and McKinney, N. P. 'The measurement of speech power.' Phonetica, 1961, 7, 65-84.

FIGURE CAPTIONS

Fig. 1. The durations of successive vocal responses emitted by a representative subject in Group Ia. There was no delay between the last shaping phase and the post-test in CRF. On the terminal questionnaire, this subject failed to describe the contingencies of reinforcement.

Fig. 2. The durations of successive responses emitted by a representative subject in Group Ib. There was a 24-hour delay between the last shaping phase and the post-test in CRF. On the terminal questionnaire, this subject failed to describe the contingencies of reinforcement.

Fig. 3. Mean response duration for each phase for individuals in Groups Ia (immediate post-test) and Ib (delayed post-test). The values are shown above the average of the Retention Index (R.I.) duration on the post-test (CRF 4).

Fig. 4. The durations of successive responses emitted by a representative subject in Group IIa. (Response durations greater than a criterion were reinforced). On the terminal questionnaire, this subject failed to describe the contingencies of reinforcement.

Fig. 5. The durations of successive responses emitted by a representative subject in Group IIb (response durations less than a criterion were reinforced). On the terminal questionnaire, this subject failed to describe the contingencies of reinforcement.

Fig. 6. Mean response durations for each phase for individuals in Group IIa (shaping of longer durations). Retention Index (R.I.) and Savings (Sav.) values are shown above the appropriate phases.

Fig. 7. Mean response durations for each phase for individuals in Group IIb (shaping of shorter durations). Retention Index (R.I.) and Savings (Sav.) values are shown above the appropriate phases.

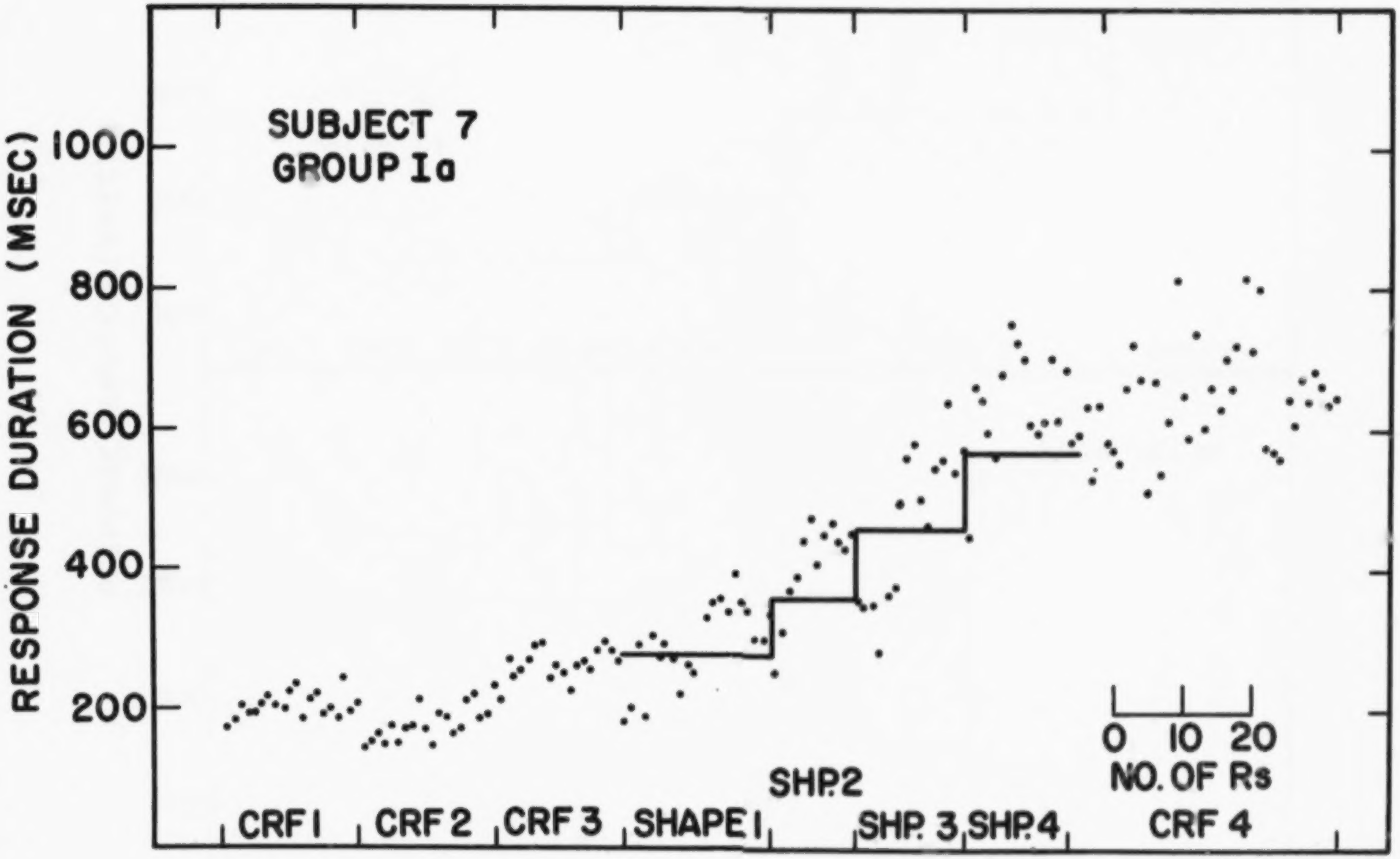
SUBJECT 7
GROUP Ia
RESPONSE DURATION (MSEC)
1000
800
600
400
200
0 10 20
NO. OF Rs
CRF 1
CRF 2
CRF 3
SHAPE 1
SHP. 2
SHP. 3
SHP. 4
CRF 4

Fig. 1

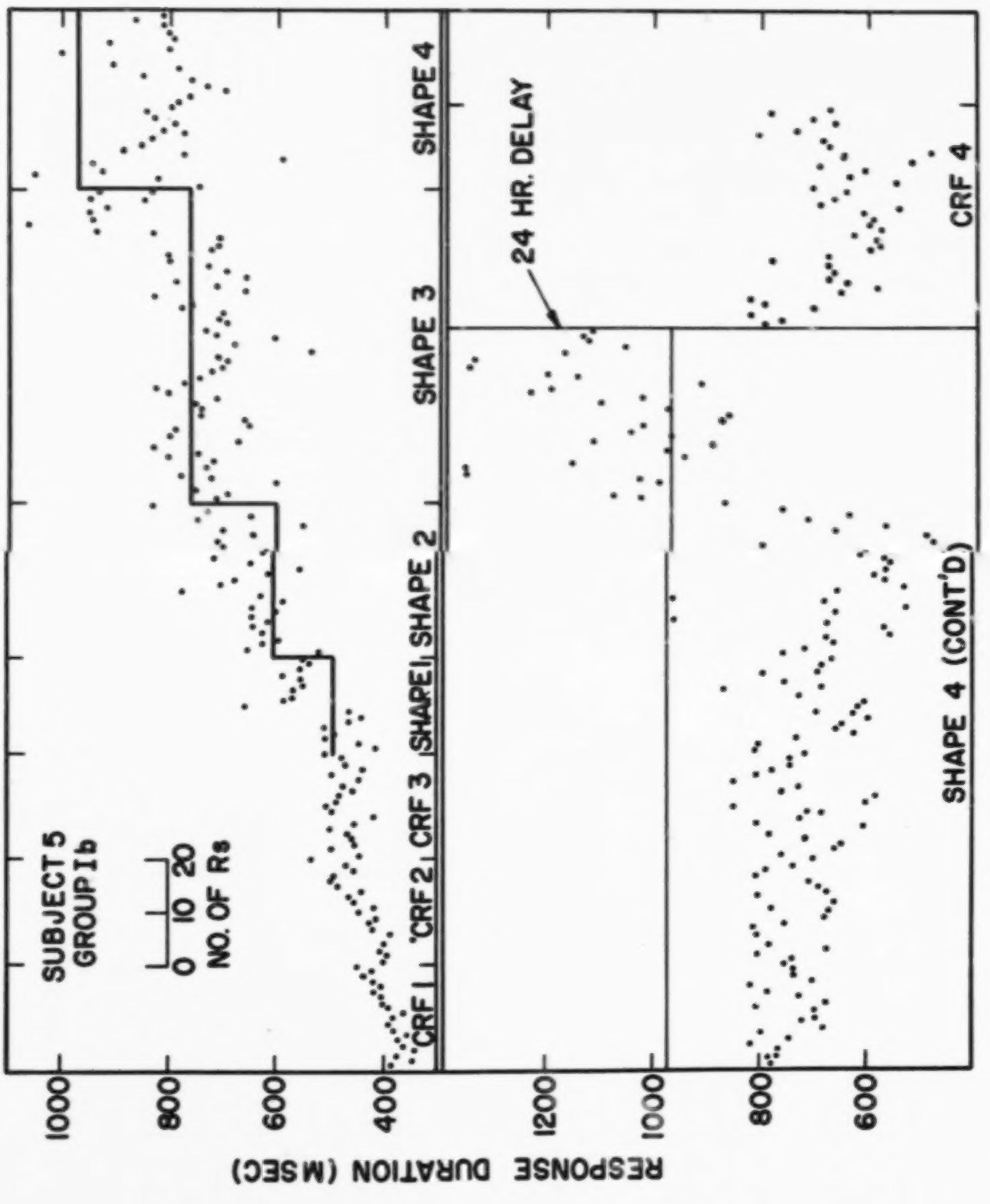
SUBJECT 5
GROUP Ib
0 10 20
NO. OF Rs
CRF 1
CRF 2
CRF 3
SHAPE 1
SHAPE 2
SHAPE 3
SHAPE 4
24 HR. DELAY
CRF 4
SHAPE 4 (CONT'D)
RESPONSE DURATION (MSEC)
1000
800
600
400
1200
1000
800
600

Fig. 2

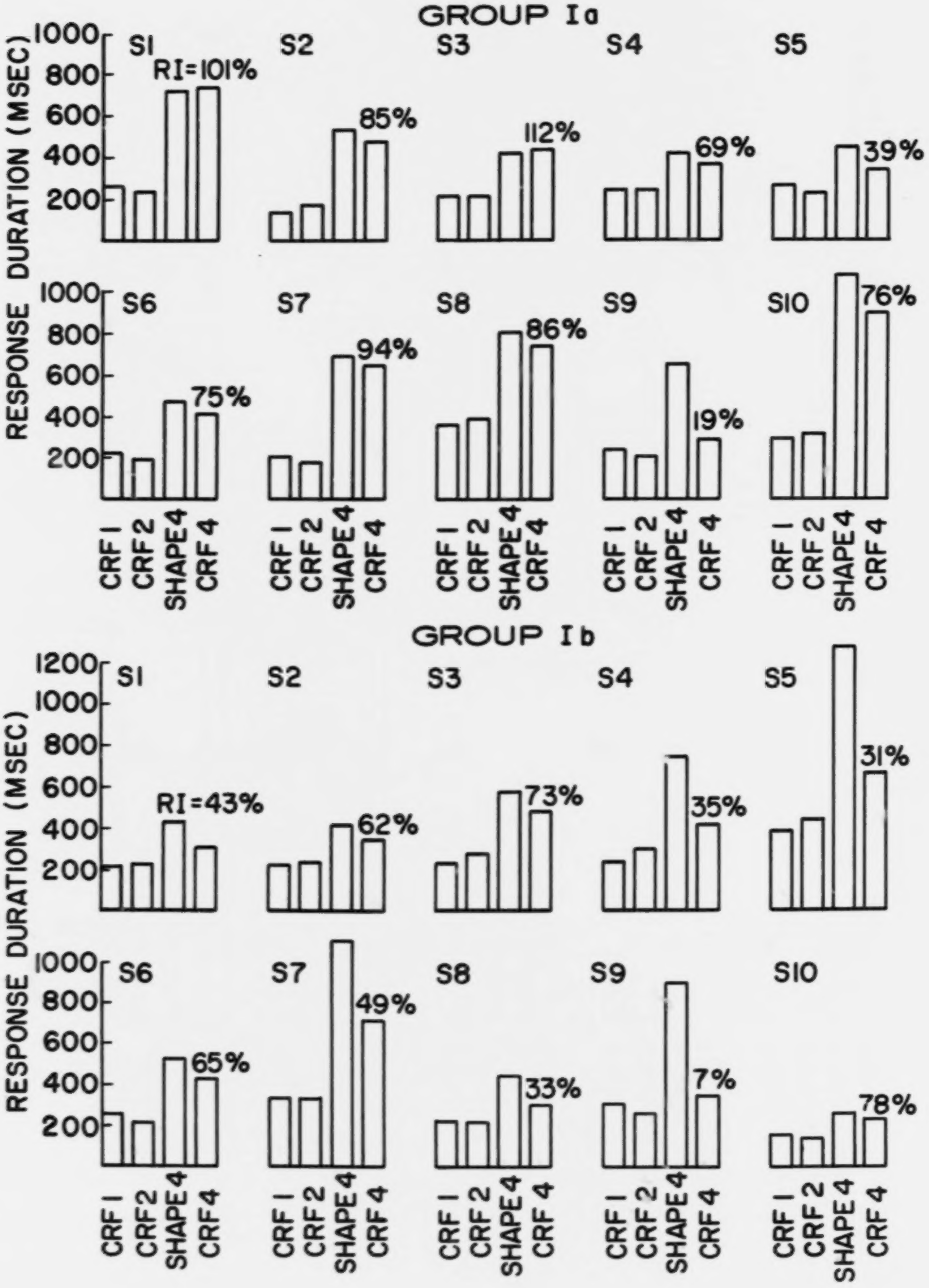
GROUP Ia
RESPONSE DURATION (MSEC)
1000
800
600
400
200
S1
RI=101%
S2
85%
S3
112%
S4
69%
S5
39%
S6
75%
S7
94%
S8
86%
S9
19%
S10
76%
CRF 1
CRF 2
SHAPE 4
CRF 4
GROUP Ib
1200
S1
RI=43%
S2
62%
S3
73%
S4
35%
S5
31%
S6
65%
S7
49%
S8
33%
S9
7%
S10
78%

Fig. 3

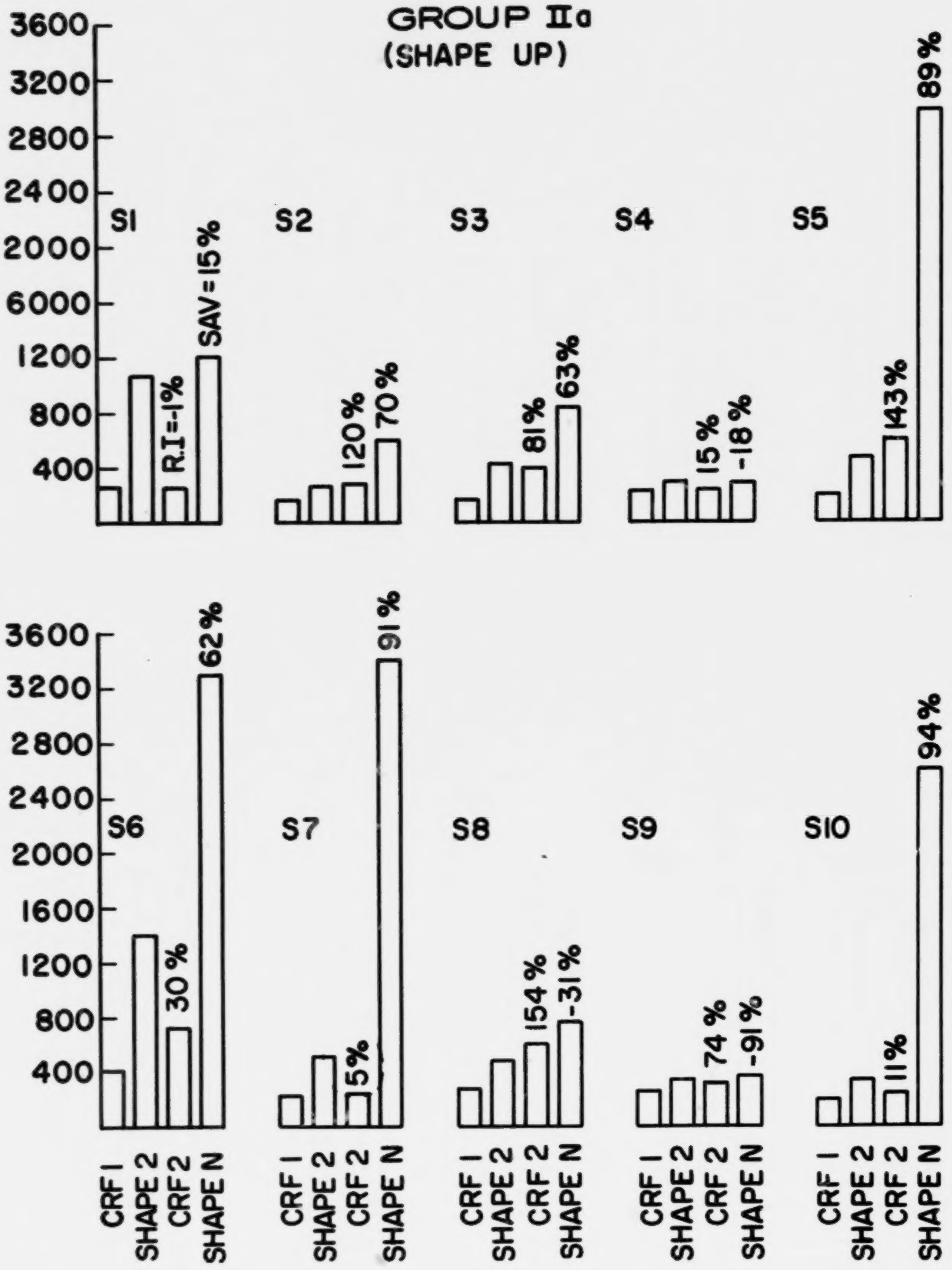
GROUP IIa
(SHAPE UP)
3600
3200
2800
2400
2000
6000
1200
800
400
S1
R.I=-1%
SAV=15%
S2
120%
70%
S3
81%
63%
S4
15%
-18%
S5
143%
89%
3600
3200
2800
2400
2000
1600
1200
800
400
S6
30%
62%
S7
5%
91%
S8
154%
-31%
S9
74%
-91%
S10
11%
94%
CRF 1
SHAPE 2
CRF 2
SHAPE N

Fig. 4

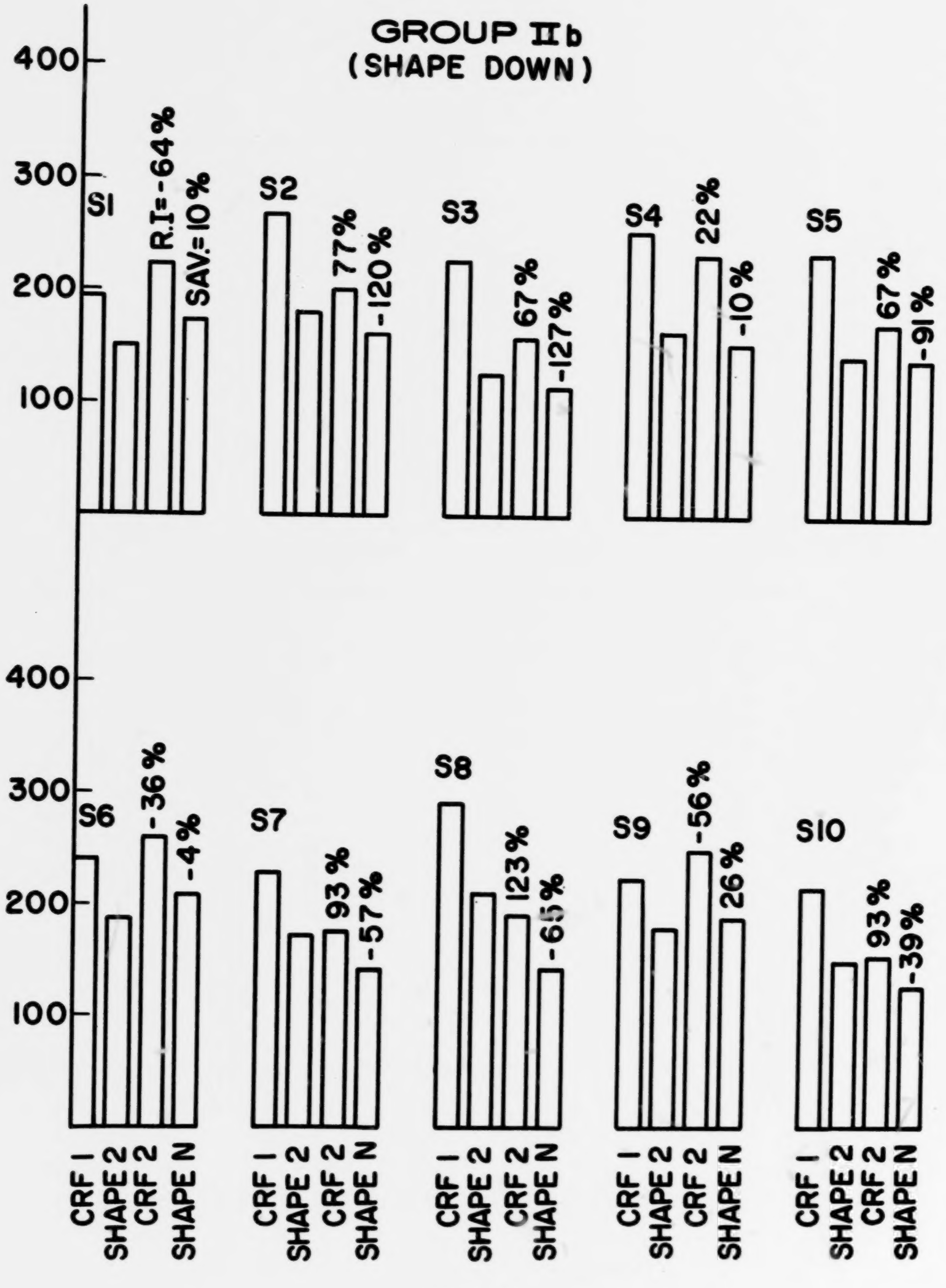
GROUP IIb
(SHAPE DOWN)
400
300
200
100
S1
R.I=-64%
SAV=10%
S2
77%
-120%
S3
67%
-127%
S4
22%
-10%
S5
67%
-91%
400
300
200
100
S6
-36%
-4%
S7
93%
-57%
S8
123%
-65%
S9
-56%
26%
S10
93%
-39%
CRF 1
SHAPE 2
CRF 2
SHAPE N
CRF 1
SHAPE 2
CRF 2
SHAPE N
CRF 1
SHAPE 2
CRF 2
SHAPE N
CRF 1
SHAPE 2
CRF 2
SHAPE N
CRF 1
SHAPE 2
CRF 2
SHAPE N

Fig. 5

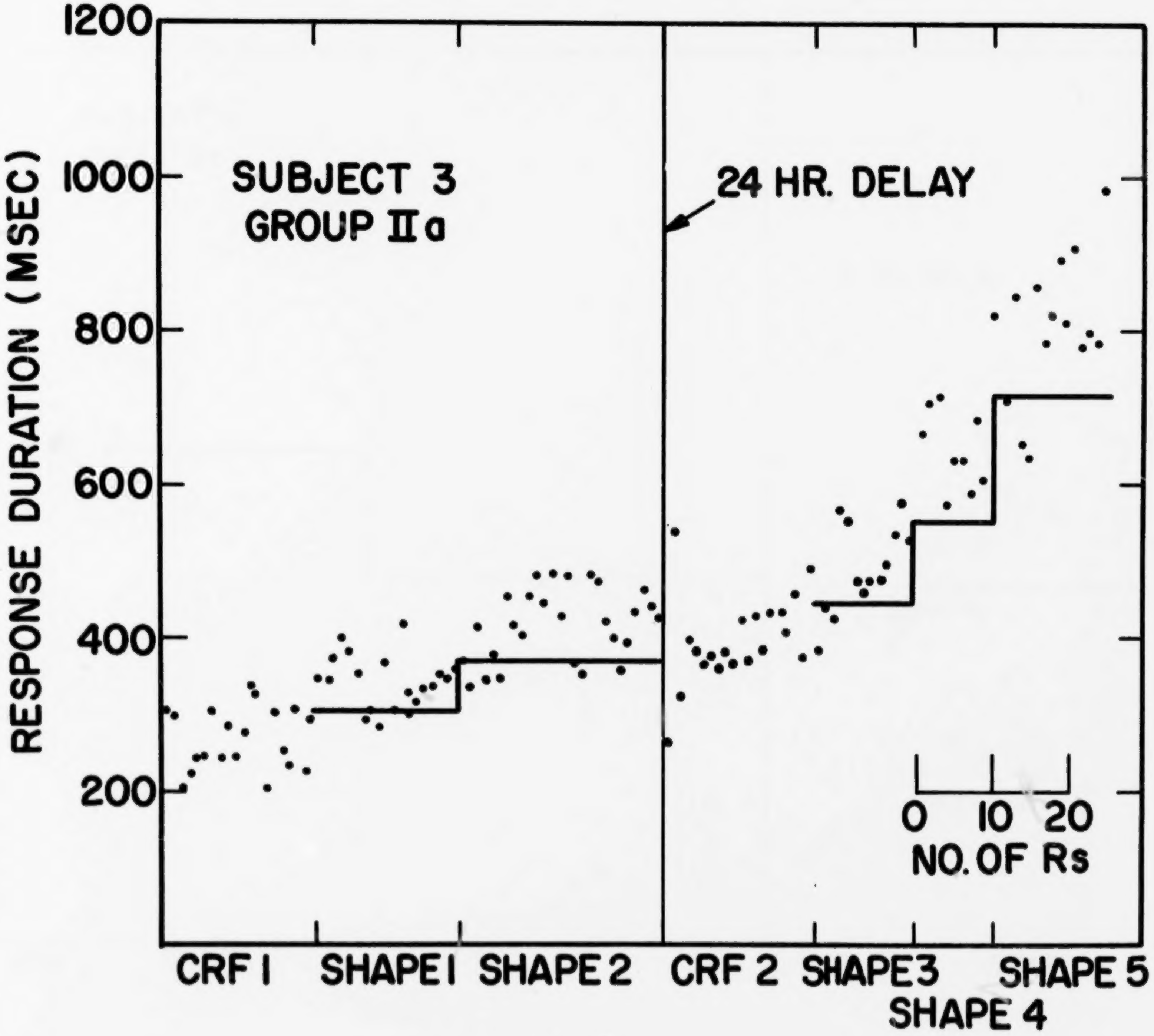
SUBJECT 3
GROUP IIa
24 HR. DELAY
RESPONSE DURATION (MSEC)
1200
1000
800
600
400
200
0 10 20
NO. OF Rs
CRF I
SHAPE I
SHAPE 2
CRF 2
SHAPE 3
SHAPE 4
SHAPE 5

Fig. 6

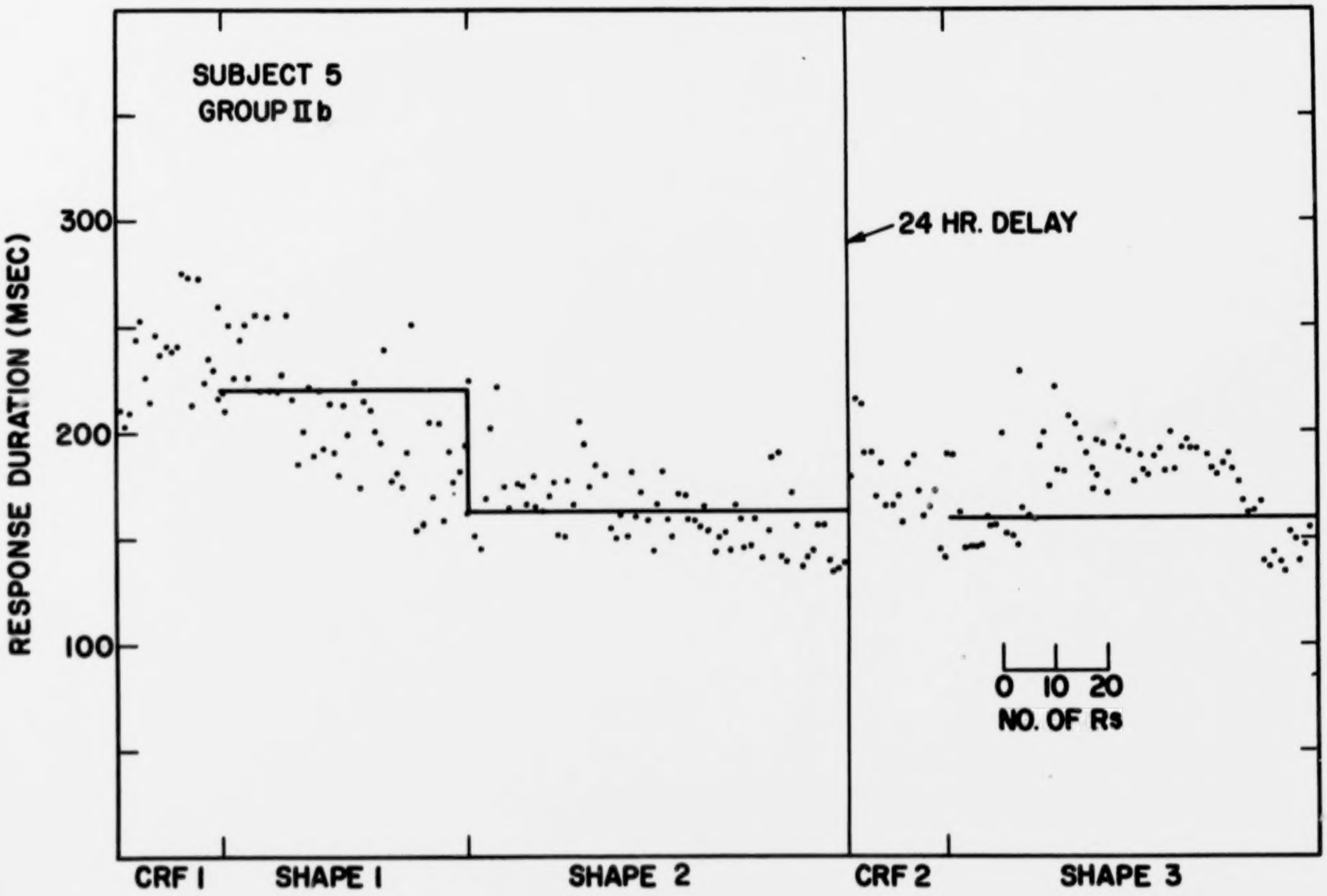
SUBJECT 5
GROUP II b
24 HR. DELAY
RESPONSE DURATION (MSEC)
300
200
100
0 10 20
NO. OF Rs
CRF I
SHAPE I
SHAPE 2
CRF 2
SHAPE 3

Fig. 7

MULTIDIMENSIONAL STIMULUS CONTROL OF THE DISCRIMINATIVE RESPONSE IN EXPERIMENTAL CONDITIONING AND PSYCHOPHYSICS

D. V. Cross

Multidimensional Stimulus Control of the Discriminative Response in Experimental Conditioning and Psychophysics

D. V. Cross

Aspects of Stimulus Control: Generalization and Discrimination

An animal that learns to emit a particular response in the presence of a specific combination of stimulus properties will emit the same response when one or more of these properties are altered or varied along underlying continua. Thus, a rat that learns to avoid electric shock by jumping over a barrier upon presentation of a 750 cps tone will also jump when a 2,500 cps tone is presented, even though it had never been paired with shock. Or the key-pecking behavior of a pigeon that is reinforced only when the response key is trans-illuminated by a red light, and is extinguished when the key is dark, may also be evoked when a green light illuminates the key. A response is said to be under stimulus control when its rate or probability of emission is observed to vary with the presence and absence of a stimulus. As exemplified above, when stimulus control is established experimentally, the fact of control is not restricted to the specific combination of physical stimulus properties initially correlated with reinforced responding, but is diffused over a wide class of specifiable stimuli. This lack of specificity that characterizes the discriminative (response-evoking) function of a stimulus is called _stimulus generalization_.

Generalization is exhibited in the behavior of the monkey that avoids all tigers, not just the first one with which it had an unpleasant encounter; and of the child who learns to call its father "Daddy" and then calls all men "Daddy." Objects viewed under different light conditions or from different angles of perspective may present widely disparate conditions of stimulation yet evoke the same previously learned identification response. In a perpetually changing environment in which stimulus conditions seldom recur in identical form, it is through generalization that an organism's behavior exhibits consistency and stability.

On the other hand, behavior may show remarkable specificity with respect to stimuli. Animals can learn to respond differentially in the presence of very slight stimulus differences. Pigeons can be trained to discriminate monochromatic lights differing by as little as one millimicron in wavelength (Hanson, 1956) The pre-medical student who is initially unable to discern even the gross features of a tissue preparation viewed through a microscope may become, through extensive training, the pathologist who is able to detect the slightest abnormalities in cell structure that indicate the presence of malignancy in a tumor. When differential responding develops in the face of generalization we call it stimulus discrimination. The professional wine-taster can make very fine discriminations; he shows little generalization among various wines. The linguist discriminates speech sounds which seem identical to (or are generalized by) most untrained listeners.

The operations entering the above definitions of generalization and discrimination serve,at one level of analysis, as dichotomous classifiers of stimuli with respect to some arbitrarily specified behavior. Two different stimuli may evoke either the same behavior (generalization), and thus be classified as equivalent with respect to the class of responses they control, or they may evoke different behavior (discrimination) and be assigned to distinct classes. The set of all experimentally specified stimulus events are thus partitioned into two classes, with class membership contingent upon the response-evoking (discriminative) function of the stimulus. This analysis by classification is not restricted to the occurrence versus non-occurrence of a single unit of behavior. Consider as an example of more complex classification an experiment in which a human subject is instructed to label by color name each of a large set of Munsell color patches. Since there are fewer commonly-used labels than there are samples in the Munsell system, the subject's behavior will partition the stimulus set into response-specific subsets or classes. Generalization with respect to the verbal response "blue" is evidenced by the large number of different color patches assigned that label, whereas discrimination is evidenced when two stimuli are assigned different labels.

In quantitative analyses of stimulus-response relations generalization and discrimination are viewed not as dichotomous aspects of stimulus control but rather as quantitative variables representing the amount or degree of discriminative control a stimulus exerts over behavior. A stimulus which consistently evokes a particular response is said to have greater control over that response than does a stimulus which evokes the response only part of the time, that is, with smaller probability. Or two different stimuli may evoke the same response but the latency of response (the time lapsing from stimulus onset to occurrence of the response) may be short for one stimulus and long for the other. A pigeon may key-peck at a steady and continuous rate in the presence of one stimulus, whereas key-pecking in the presence of a generalized stimulus may be characterized by breaks in rate, that is, by periods of responding interspersed with not responding. Measures of response strength derived from these observations (or from other experimental operations) provide a basis for differentiating stimuli which are otherwise given the same classification. The stimulus generalization gradient--defined operationally in relation to changes in the discriminative stimulus as a succession of decrements in response probability or in some other measure of response strength--identifies the class of stimuli that are effective in evoking a particular response and, in addition, defines the degree of membership for each stimulus in that class.

From a strictly operational point of view generalization and discrimination may be considered alternative and opposite ways of describing the same set of stimulus-response relations (a view either explicitly or implicitly held by numerous investigators: e.g., Lashley & Wade, 1946; Gibson, 1959; Mednick and Freedman, 1960; Prokasy & Hall, 1963; Brown, 1965). "The experimentally conditioned [stimulus], and any other stimulus on a physical-property continuum with it, are said to generalize to the extent that they overlap in capacity to evoke R; to the extent that they differ in evocative capacity, they are said to be discriminated." (Schoenfeld & Cumming, 1963, p. 225) If two stimuli set the occasion for different levels of responding (represented by different points on the generalization gradient) they are exerting differential control over behavior (discrimination) and the difference in their response measures may be interpreted either as a measure of discrimination or as a measure of generalization difference.

Rules of Combination in Multidimensional Stimulus Control

Stimulus generalization and discrimination have a long and interesting history of experimental study and they have been investigated from many viewpoints, but little attention has been directed toward the problem of how generalization (or discrimination) occurs in the presence of stimuli that vary simultaneously on two or more dimensions as compared with generalization (or discrimination) occurring in the presence of stimuli that vary along single dimensions only. *What rule of combination predicts the behavioral control exerted by a multidimensional stimulus as a function of the control exerted* *separately by each of its component properties*? This problem pervades all levels of behavior, simple and complex, animal and human, since the stimuli which set the occasion for differential responding outside of the laboratory characteristically differ on two or more dimensions and may even involve different sensory modalities. How does the behavior evoked in these situations compare with discriminative behavior that is under the control of unidimensional stimuli? What are the rules that govern how unidimensional stimulus effects combine to produce multidimensional stimulus control? These questions may be raised with regard to a variety of experimental problems: whether we are concerned with the course of development of stimulus control, as in discrimination-learning experiments; with the specificity of stimulus control as in generalization experiments; with the relation between specific properties of stimuli and presumptive measures of sensory magnitude, as in psychophysical experiments; or with the general question considered in multidimensional scaling theory of how different stimulus attributes combine to affect the total similarity of perceived objects.

Experimental Conditioning with Compound Stimuli

Research in the area of experimental conditioning has been concerned with the operations that establish a stimulus-response relation where none, or a different one, existed before, usually with emphasis on the course of development or the specificity of control as a function of other variables (e.g., drive level, reinforcement schedule, etc.). These experiments include studies of stimulus generalization, in which a response is brought under the control of a particular stimulus complex and then tested in the presence of other stimuli sampled from a common physical-property continuum. Included also are conditioning experiments and discrimination-learning experiments in which the dependent variable is either the different number of stimulus presentations or trials required to establish control, or the degree of control established as measured by differential amounts of conditioning or the number of correct discriminations obtained. Few of these studies have investigated conditioning as a function of the dimensionality of the stimulus variable.

The studies reviewed here were not concerned with the specific rule of combination that predicts multidimensional stimulus control but rather with the more elementary question of whether stimuli differing on two or more dimensions exercise more behavioral control than do stimuli which differ on one dimension only. In spite of the quantitative limitations, the results of these studies provide a basis for some rather broad inferences about rules of combination in the multidimensional case. For example, Miller (1939) found that a combined auditory and visual stimulus (a drop in pitch of a continuous tone, and the movement of a pointer), paired with electric shock to the cheek, resulted in faster and stronger conditioning of an eyelid response in humans than did either stimulus property used alone. Similarly, Eninger (1952) found that rats formed a discrimination much more rapidly when auditory and visual cues were combined (tone vs. no-tone, and white vs. black) than when either of these sets of cues was presented alone. In a generalization study, Fink & Patton (1953) controlled certain visual, auditory, and tactile characteristics of the stimulus situation in which rats learned to drink water from a tube. After establishing a stable baseline of drinking rate, the investigators altered one or more of the stimulus properties under their control. They found that any change led to a decrement in drinking: light changes were most effective, sound changes were moderately effective, and tactual changes were least effective in modifying the strength of the learned drinking response. More important, however, they found that a change in any two stimulus properties resulted in a greater response decrement than a change in any single property; when all three stimuli were changed, maximum response decrement occurred. White (1958) trained children to pull a handle to get marbles, using a card of a certain hue and saturation (Munsell 10GY 8/6) as a discriminative stimulus. Then, for different groups of subjects, he changed the hue of the stimulus card, its saturation, or both, in a series of unreinforced test trials. There was greater generalization decrement to test stimuli which differed from the training stimulus in both dimensions than to those which differed in either of the dimensions alone. That is, the properties of hue and saturation combined to exert more control over differential responding than either property exerted alone. A generalization suggested by these results is that every stimulus property that happens to be correlated with the contingencies of reinforcement requiring differential responding acquires some degree of stimulus control; when stimulus properties are presented in combination, their effects combine in some manner to exert an even greater degree of control.

On the other hand, it has been found, in some discrimination experiments, that stimuli differing in two properties are discriminated with no greater ease or efficiency than stimuli differing *in the single one* of these properties which is most readily discriminated. In other words, the degree of behavioral control exerted by the several properties of a compound stimulus is no greater than that exercised by the component property which, by itself, is most effective. Harlow (1945) trained monkeys to respond discriminatively to stimuli differing in either color, form, or both. Stimuli differing in both color and form, or in color alone, were discriminated equally well, although both were discriminated more readily than form. This result was also found by Warren (1953) who tested monkeys on discrimination problems involving stimuli which differed in color, form, size or in the four combinations of these properties taken two or three at a time. Stimuli which differed in color and some other property were little, if any, more discriminable than those differing in color alone. In a later experiment, Warren (1954) trained monkeys to discriminate pairs of stimuli which differed simultaneously in color, form, and size. After establishing stimulus control, the number of discriminative cues was reduced by eliminating stimulus differences in one or two dimensions. In each instance, performance scores were depressed. Discriminative control was almost completely lost when color differences were eliminated, regardless of which other property or combination of properties remained. Some control was lost when form and size differences were both eliminated, and the smallest decrements in performance occurred when only form or size differences were eliminated, size being the least effective.

The experimental results considered first suggest a distribution of behavioral control over the several dimensions of a complex discriminative stimulus, while the others suggest that a single dimension overshadows the rest and is dominant in controlling differential responding. In other words, the combination of stimulus properties results not in a summation of their separate effects but in an effect equal to that of the property which, alone, is most effective. It is evident that no one combination rule can account for these contradictory findings.

Metric Analyses in Multidimensional Psychophysics. Traditionally the study of "similarity" relations imposed by behavioral measures on pairs of stimuli that differ with respect to several physical properties at once has been called "multidimensional psychophysics." Analysis of experimental data obtained in this line of research has almost universally been guided by the assumption that the collection of N stimuli employed in an experiment can be represented by a spatial configuration of "points" in a multidimensional Euclidean space. The configuration is determined by the N^2 inter-point distances which are derived from behavioral measures obtained on each stimulus pair. The experimental data can be displayed in an N x N matrix, with each row and corresponding column associated with a particular stimulus and each cell containing a response measure, P_{ij} (pertaining to the row stimulus S_i and the column stimulus S_j), which depends on the experiment and behavior sampled. These are called proximity measures by Shepard (1962) because they are interpreted as indicating how closely stimuli S_i and S_j are related. In a generalization experiment P_{ij} may represent the amount of generalization between a training stimulus S_i and a test stimulus S_j, as measured by response probability, latency or rate; or P_{ij} may correspond to the percent "same-different" judgments evoked in a discrimination task; or to the "similarity" of S_i and S_j as rated by subjects on a scale from, say, 0 to 10. These are called symmetric or conditional proximity data by Coombs (1964), depending on whether $P_{ij} = P_{ji}$, or $P_{ij} \neq P_{ji}$.

Whatever the behavioral measures obtained in a given experiment a function is postulated which transforms them into a set of presumptive distances D_{ij} which may be analyzed by a multidimensional scaling algorithm, first proposed by Young and Householder (1938), and subsequently refined by Torgerson (1952), and by Messick and Akelson (1956). The purpose of the analysis is to determine the effective dimensionality of the space in which the stimuli are assumed to be imbedded and to determine the projections of the stimuli on each of the dimensions involved. The method invokes the law of cosines in order to transform the matrix of distances $[D_{ij}]$ into a matrix $[B_{ij}]$ of scalar products of stimulus vectors, having as their common origin the centroid of the points. The matrix $[B_{ij}]$ is then factor-analyzed in much the same way as correlation coefficients. The coordinates for each stimulus on orthogonal axes in psychological space are then given by their factor loadings.

A study by Shepard (1958) will serve to illustrate this procedure. The study was an identification-learning experiment in which subjects were required to learn a different arbitrarily-specified response to each of nine Munsell colors, varying in brightness (value) and saturation (chroma). In the course of the experiment generalization occurred, that is, the response appropriate to, say, stimulus S_i was also evoked by S_j. Generalization from S_i to S_j, denoted by $g(i,j)$, is represented by the ratio $g(i,j) = n_{ij}/n_{ii}$ where n_{ij} is the number of S_j presentations resulting in the response appropriate to S_i, and n_{ii} is the number of times the response is correctly evoked by S_i. Shepard postulated, after Hull (1943), that $g(i,j)$ is an exponential decay function of the psychological distance between S_i and S_j, that is,

$$g(i,j) = e^{-kD_{ij}} \qquad \text{or} \qquad D_{ij} = -k' \log_e g(i,j) \ .$$

Thus, the distance between the points, in "psychological space", corresponding to S_i and S_j is given by the logarithm of the corresponding measure of generalization. However, two separate measures of generalization were obtained for each stimulus pair, $g(i,j)$ and $g(j,i)$, and each gave rise to a different estimate of the same distance. So, in order to obtain a symmetric distance-matrix, Shepard simply averaged the two. The resulting distances were then reduced to an orthogonal set of coordinates by the method indicated above. Actually, this is a somewhat simplified account. Shepard also calculated weights for each stimulus, which were intended to correct for response biases and the asymmetries that occurred, and determined a constant which theoretically took account of the essentially random responses occurring during early learning trials. In an attempt to evaluate the goodness-of-fit of the scaling model to the empirical data, Shepard reconstructed response measures from the weights and final factor loadings computed for each stimulus. He reported that approximately 99 percent of the variance in the original response measures was accounted for by reconstruction.

Two important theoretical assumptions guided this analysis. The first concerns the form of the generalization function, and the second concerns the rule that governs how differences along two or more dimensions combine to determine the overall difference between two stimuli. The assumption that psychological space is Euclidean in character requires that the total difference between two stimuli be defined according to the Pythagorean theorem, as the square root of the sum of squared differences in projections upon orthogonal reference axes (dimensions). This is the combination rule adopted explicitly, or implied, by most multidimensional scaling models. On the face of it there appears to be no good reason to believe that this is in fact how stimulus control is distributed over the various dimensions of psychological space. There is certainly no conclusive experimental evidence to indicate that this is so. The truth of the matter is that the adoption of the Euclidean combination rule has simply followed the dictates of mathematical convenience. The analytic procedures of most scaling models exploit certain mathematical relations that are valid only in Euclidean spaces.

The methods of multidimensional scaling have been and will continue to be useful techniques for: (1) the construction of coordinate systems descriptive of stimuli of unknown dimensionality or of stimuli on which the physical measurements that can be made are not directly related to the behavior under investigation; or (2), for reducing a large set of correlated measures to a smaller set of orthogonal linear components which account for most of the variance in the original measures. However, the practice of demanding reference to a Euclidean metric for behavioral data that come equipped with their own metric (the combination rule that actually generates the data) imposes constraints upon the analysis that may result in an unwanted misrepresentation of the behavior, in terms of both the combination rule and the dimensionality of the solution. If the selection of a particular metric embedding is to have any heuristic value it should provide not only a convenient multidimensional representation of the stimuli, but also a valid description of the behavior in question, and a basis for important theoretical conclusions and techniques for comparing and testing alternative theories of behavior.

Unfortunately, very few studies have approached the problem of multidimensional data analysis from the point of view of evaluating the metric, or combination rule, that actually generated the data. Several authors have suggested that a non-Euclidean metric may be appropriate (e.g., Attneave, 1950; Galanter, 1956; Torgerson, 1958; Shepard, 1964; Coombs, 1964). Coombs (1964) has discussed the advantages of exploring alternative metric representations of behavioral data. In his own words "one of the most desirable consequences of developing alternative models and their algorithms for data analysis is that *their existence destroys any naive complacency with any one model and leads to a search for ways of testing and comparing alternative theories*." (p. 206, italics in original) The point is that the algorithms used in multivariate data analysis and the metric embeddings of multidimensional scaling models are, by their nature, theories about behavior. They are theories about how unidimensionally-scaled stimulus-effects combine to determine the behavior evoked in multidimensional situations. There is not necessarily just one combination rule to which all behavior conforms, but rather alternative rules which may correspond to alternative theories about behavior, or to alternative behaviors, or to behavior under different circumstances, or under the control of different kinds of stimuli.

Metrization of Multidimensional Stimulus Control

If the properties of discriminative behavior under multidimensional stimulus control can be suitably represented by the mathematical properties of *metric spaces*, they can then be analyzed in terms of the mathematical procedures of the geometries associated with these spaces. This would not only assist us in the practical business of dealing with multidimensional stimulus variables, but also provide a formal basis for characterizing and contrasting alternative theories or hypotheses about behavior.

The basic concept underlying the notion of a metric space is that of the distance between two points. The properties of distance will depend to some extent on the space considered; but certain basic properties are definitive and assumed always to hold. These are: (1) the distance between any two points is non-negative and only the distance of a point from itself is zero; (2) the distance between any two points is symmetric, that is, the distance from the first point to the second is the same as that of the second from the first; and (3) for any three points the distance between one pair is no greater than the sum of the other two distances, a condition called triangle inequality. These assumptions can be stated a little more exactly by formulae. We can consider a set S such that to each pair x,y of its elements we can associate a non-negative real number $d(x,y)$, called their distance, which satisfies the conditions: (1) $d(x,y) = 0$, if, and only if, $x = y$; (2) $d(x,y) = d(y,x)$; (3) $d(x,z) \leqq d(x,y) + d(y,z)$, for all x,y,z in S. The function $d(x,y)$ is said to be a metric for S. Any set (or space) S is said to be metrizable, if a distance between its elements (or points) can be defined. The ordinary physical space of three dimensions is a metric space with Euclidean distance. The above conditions are quite general, however, and are also satisfied by the following example. Take a set of color patches and for each pair x,y define the distance $d(x,y) = 0$, if both members of the pair are labelled with the same color name, and define $d(x,y) = 1$, if labelled differently. The function $d(x,y)$ could serve as a metric for the color domain, but it does not provide a basis for a very interesting or especially informative model for color perception.

The metric axioms clarify the conditions under which measures of stimulus control may be treated as measures of distance. However, in the collection of single stimulus data, as in generalization and some discrimination training paradigms, the task of empirically verifying each of these conditions for all stimulus pairs can be formidable. Consider a stimulus generalization decrement, defined as the difference in the response measure produced by the training stimulus and each test stimulus. Direct tests of the symmetry and triangle inequality properties of distance requires N replications of the study, each time with a different training stimulus, in order to fill in the N^2 cells of the data matrix. This undertaking can be expensive in terms of research time and in the consequences of failure to control all variables affecting stimulus control in the separate replications.

On the other hand, if the effective dimensions of stimulus control are presumed to be known (or at least referable to the physical dimensions of the stimulus), then the question of whether behavioral measures define a metric for the stimulus set can be answered by evaluating empirically the combination rule that relates the amount of control exerted separately by each unidimensional stimulus change to the total control exerted when all changes are effected simultaneously.

Metric Implications of Equal-level Surfaces. Let $P_1, \ldots, P_n$, respectively, denote each of the n physical-property continua or dimensions (e.g., wavelength, intensity, size, etc.) called upon to operationally characterize the stimuli involved in, say, a generalization experiment. Each stimulus can be represented by a point in a physically-defined space P, where

$$P \equiv P_1 \times \cdots \times P_n \equiv \{(p_1, \ldots, p_n) : p_i \in P_i \ (i = 1, \ldots, n)\}.$$

The space P may be viewed as an n-dimensional Cartesian space with orthogonal reference axes P_i $(i = 1, \ldots, n)$ intersecting at a common point θ which is taken as the origin of the space. If we take as θ the point representing the particular combination of physical aspects that specifies the initial training stimulus in the generalization experiment, then each test stimulus p may be represented by an ordered n-tuple $p = (p_1, \ldots, p_n)$, the components of which pertain to the signed physical differences between p and θ on each of the n dimensions of P; the quantities p_i $(i = 1, \ldots, n)$ are the relative coordinates of p with regard to θ.

Each P_i is a unidimensional metric space with the metric defined by $d_i(p_i, q_i) = |p_i - q_i|$ for every pair of points $p_i, q_i \in P_i$; in other words, $d_i(p_i, q_i)$ is the physical separation of $p, q \in P$ on the single dimension P_i. Since the Cartesian product of n metric spaces is also a metric space, it is possible to define a function $d: P \times P \rightarrow R$ which assigns to each pair of points $p, q \in P$ a real number $r \in R$ (the set of real numbers) interpreted abstractly as the physical distance between the stimuli denoted by p and q. We might ask the following question: under what conditions is it possible to construct a metric for P in such a way that the distance $d(\theta, p)$, defined for fixed θ and any p in P, is equivalent to the difference in behavioral control (e.g., generalization decrement) exerted by the two stimuli represented by θ and p.

Let $\overline{g}(\theta, p)$, or simple $\overline{g}(p)$, represent the difference in behavioral control exerted by θ and p in stimulus generalization, that is, $\overline{g}(p) = g(\theta) - g(p)$, where g is the generalization function which we assume has its maximum valuation at θ. Consider the point set $U \equiv \{p \in P: \overline{g}(p) \leq u\}$, containing only those points in P that produce a generalization decrement relative to θ less than or equal to some arbitrary positive value u. The measure of generalization decrement $\overline{g}(p)$ can be taken as a metric for P if and only if:

(i) the point set U is symmetric with respect to θ,

(ii) the point set U is convex, and

(iii) the boundary of U is homothetic to the boundary of every point set generated by a different valuation of the parameter u.

The condition of symmetry simply requires that, if the point p belongs to U, then the point $-p = (-p_1, \ldots, -p_n)$ also belongs to U. The convexity condition requires that for each pair of points p, q ε U the line segment joining them is entirely in U; in other words, each point $r = \lambda p + (1 - \lambda)q \equiv (\lambda p_1 + (1 - \lambda)q_1, \ldots, \lambda p_n + (1 - \lambda)q_n)$, $0 \leq \lambda \leq 1$, is also in U.

The boundary of U, which we will denote by U', is the set of all points p for which $\bar{g}(p) = u$. If (i) and (ii) hold, then this set constitutes a centrally symmetric, convex surface in P. Condition (iii) requires that all such surfaces in P be of the same shape, that is, encased one into the other and produced through dilations or contractions in the ration u:1, where u is a real positive number.

For u = 1 the set U' is called the standard unit surface for P, and it enters the definition of a distance function for P in the following way: for any two distinct points p, q ε P, let the ray with initial point θ, which is parallel to the directed line segment $\overrightarrow{pq}$ (and with the same sense), meet the surface U' at point r. The Minkowski distance of the points p and q, denoted m(p, q), is defined in terms of the Euclidean metric associated with P in the following way:

$$m(p, q) = \frac{e(p, q)}{e(\theta, r)}$$

where e(p, q) and $e(\theta, r)$ are the Euclidean distances in P of the points p,q and θ, r respectively.

If $\bar{g}(p) = t$, then $t = m(\theta, p) = e(\theta, p)/e(\theta, p')$, where p' is the intersection of the surface U' with the ray (θ, p), and t is the ratio of similitude of a homothetic transformation of P, which carries U' into a surface passing through P.

For the points on any Euclidean line L in P, the change in Minkowski metric is merely a change of scale; the distances $m(p, q)$ and $e(p, q)$ are proportional on each line L.

A Minkowski metric in P is invariant under translations and central reflections of the coordinate axes of P. It is not, however, invariant under rotations of the axes about the origin. The Euclidean metric is the only specialization of Minkowski distance that has rotation invariance.

Two Minkowski metrics $M_a(p, q)$ and $m_b(p, q)$ in P are isometric if and only if an affinity exists which maps the unit surface U_a': $m_a(\theta, p) = 1$ on the surface U_b': $m_b(\theta, p) = 1$. Special cases of such affine transformations are rotations of the coordinate system, homothetic transformations (a uniform stretching or shrinking of the coordinate axes) and simple elongations and compressions of axes, sometimes called one-dimensional strains. A Minkowski metric in P is isometric to a Euclidean metrization of p' (where P' results from an affine transformation on the points of P) if and only if the unit surface of P is an ellipsoid.

More complete discussions of the mathematical properties of Minkowski spaces can be found in Hancock (1939) and Busemann (1950, 1955).

It is readily seen that if $\overline{g}(p)$ represents a Minkowski metric for P, then $\overline{g}$ must be a positive definite, convex and homogeneous function of the first order in its n variables $p_1, \ldots, p_n$. That is:

(1) $\overline{g}(p) > 0$ and only $\overline{g}(\theta) = 0$

(2) $\overline{g}(tp) \equiv \overline{g}(tp_1, \ldots, tp_n) = |t| \; \overline{g}(p)$ for any real number t

(3) $\overline{g}(p + q) \equiv \overline{g}(p_1 + q_1, \ldots, p_n + q_n) \leq \overline{g}(p) + \overline{g}(q)$.

In this form the Minkowski distance of p and q in P is defined as $m(p,q) = \overline{g}(p - q) \equiv \overline{g}(p + (-1)q) \equiv \overline{g}(p_1 - q_1, \ldots, p_n - q_n)$.

Conversely, any combination rule or function F that is defined in P and has these properties is a Minkowski metric. It is evident that the metric axioms stated previously are satisfied by the function $F(p - q)$. The symmetry of distance follows from (2) by setting $t = -1$. From the convexity property (3) the law of triangle inequality follows, for if we set $p' = p - r$ and $q' = r - q$ then

$$F(p' + q') \leq F(p') + F(q')$$

$$F(p - q) \leq F(p - r) + F(r - q)$$

or $$m(p, q) \leq m(p, r) + m(r, q).$$

Treatment of Nonlinear Stimulus-Response Relations. Shepard (1964) was the first to propose that the combination rules governing multidimensional stimulus control can be determined empirically by investigating the shape of the surfaces in P, corresponding to prescribed levels of control:

> With respect to the physically defined space, a knowledge of these rules is equivalent to a knowledge of the shape of the locus of all stimuli that have any prescribed degree of similarity to any prescribed standard stimulus ... the question of this shape turns out to be equivalent to the question of just what metric is appropriate for the psychological space. (p. 56)

Shepard points out that if psychological space is Euclidean in nature, then (providing the unidimensional psychological scales involved are affine transforms of the corresponding physical continua) the isosimilarity contours in physically-defined space of two dimensions must be affinely equivalent to circles, that is, elliptical in shape. If these contours are found to have some non-elliptical form, then psychological space cannot be Euclidean.

However, if generalization decrement is not symmetrically linear with changes in the physical stimulus on each physical property continuum then neither a Euclidean nor any other Minkowskian metrization of P is possible. Unfortunately, this appears to be the rule rather than the exception in most investigations of stimulus control. Where intensive physical continua are involved, it is notoriously the case that discrimination is not an invariant function of the physical separation of two stimuli, but depends on their location on the continuum. Collaterally, the generalization decrement produced by an increase in stimulus intensity does not invariably match that produced by an equivalent decrease in intensity. Furthermore, the generalization decrement produced by a stimulus difference of two units is not invariably twice that produced by a stimulus difference of one unit. These outcomes obviously violate the property of homogeneity of metric functions, and in some instances may lead to a violation of convexity (hence triangle inequality) as well. However, nonlinearity of stimulus-response relations does not preclude the possibility of mapping P into a space that *is* metrizable in terms of stimulus response relations, and which is Minkowskian.

Three cases arise in which, respectively, transformations of the stimulus variables, transformation of the response variable, or both may be required in order to transform the coordinates of P into the coordinates of a metric space P' (defined as psychological space), wherein these variables may enter relatively simple and psychologically meaningful relations, and wherein a combination rule may obtain that has generality as a principle of behavior. The cases follow:

1. Choice of stimulus scale. The space P may not be metrizable because of nonlinear deformations of the *equal-level surfaces* of P in directions parallel to one or more reference axes. See, for example, the two-dimensional *equal-level contours* plotted in Fig. 1(a). These contours are neither centrally symmetric, homothetic, nor convex. This outcome may occur if generalization gradients are symmetrically linear when the stimuli are 'appropriately' scaled in logarithmic units, and a simple "additive" combination rule holds for generalization decrement in two dimensions. In this instance a Fechnerian (log) transformation on both axes of P maps P into P' that has level contours like those depicted in Fig. 1(d).

The empirical operations involved in the measurement and control of physical stimuli result in quantities or magnitudes, such as sound intensity, frequency, wavelength, etc., which generally enter very complex relations with behavior; on the other hand, relations defined in terms of "loudness," "pitch," or "hue" (represented by psychophysical transformations of the corresponding physical scales) are likely to be simpler and to have greater generality. By insisting that principles of stimulus control be taken at the level of the physical operations involved in specification of his experimental variables, an investigator may never discover important principles that relate behaviors evoked under different stimulus conditions.

> However much we wish it were not so, we shall not be able to make sense out of data in animal psychophysics (this includes generalization) without taking account of functions that are usually attributed to the 'sensory systems.' The transformations imposed by the sensory system must be 'subtracted out,' as it were, before other regularities will become clear. (Blough, 1965, p. 35)

Although we may not know in advance what psychophysical function is appropriate for each physical continuum, we can "subtract out" this source of nonlinearity in stimulus control by rescaling the coordinate axes of P in units of generalization decrement, thus mapping P into $P' = \{(P_1', \ldots , p_n'): p_i' \epsilon P_i' \ (i = 1, \ldots , n)\}$ where $p_i' = \text{sgn } p_i \ \overline{g}(p_i)$ and sgn p_i is equal to 1 for $p_i > 0$, 0 for $p_i = 0$, and -1 for $p_i < 0$ (recalling, here, that P_i was defined as the signed physical difference between θ and p on P_i). This transformation has an effect equivalent to that of initially scaling the stimulus continua in appropriate psychophysical units. If the level contours obtained in P are affinely equivalent to those in Fig. 1(a), then a rescaling of the axes in units of generalization decrement would result in countours in P' affinely equivalent to those in Fig. 1(d). The Minkowski distance from θ to x in Fig. 1(d) is equal to the ratio of the Euclidean lengths θx' (where x' lies on the "unit contour") and θx.

2. Choice of response scale. Rescaling the stimuli in units of behavioral control forces symmetry on each separate dimension of P' and may impose symmetry in the large, so that $\overline{g}(p') = \overline{g}(-p')$ for all $p' \varepsilon P'$. However, this outcome does not guarantee that P' will be metrizable in terms of $\overline{g}$. Consider the level contours shown in Fig. 1(b). Rescaling the axes of P in behavioral units transforms P to the space represented in Fig. 1(c) that has level contours which are centrally symmetric and convex, but not homothetic. The shape of these contours changes gradually with level of generalization decrement. This means that the behavioral measures leading to this result cannot serve as measures of distance between stimuli.

In order to succinctly express these hypothetical findings in terms of a suitable metric, it is necessary to transform the original generalization measures, that is, to redefine stimulus control in terms of an alternative response scale. In the example at hand a logarithmic transformation of the dependent variable changes the structure of P' as represented in Fig. 1(c) to the simple metric structure depicted in Fig. 1(d). In fact, this is the only transformation that will simultaneously map all level contours of Fig. 1(c) into a set of contours that have, regardless of level, the same centrally symmetric and convex shape. No alternative transformation exists, for example, that will map the contours of Fig. 1(c) into a set of concentric circles. If this outcome is observed with real data the suggested transformation need not be viewed as an arbitrary handling of data, but rather as implying that the response measures are related exponentially (in this particular example) to alternative response measures in terms of which this simple metric structure would have been obtained directly.

Recalling that generalization decrement was defined as $\overline{g}(p) = g(\theta) - g(p)$, where g is the generalization function, we must turn our attention to the measurement operations entering the experimental definition of generalization. It is clear that the form of the empirical generalization gradient and the questions of whether generalization decrement is linear or nonlinear with the size of the stimulus difference in psychophysical units, and of whether a metric combination rule applies, depends on what property of behavior is being measured, and how it is measured. Generalization can be, and has been, empirically defined in terms of a variety of response measures: response probability, number of responses emitted in a fixed interval of time, response latency, rate, amplitude, number of responses to complete extinction, etc. These variables are typically not linearly related, but they are presumed to be mutually interrelated in such a way that if a metric can be established for one, then appropriate transformations on the other measures will result in the same metrization of stimulus control. As Coombs pointed out in connection with the development of mathematical models for analyzing behavioral data:

> The problem is not to be formulated in terms only of transforming these response measures to construct psychological spaces, but rather of transforming them to construct interlocking systems which mutually relate the response measures to each other so that we would obtain the same metric space from each. The models for processing these several measures should be mutually compatible in the sense of yielding at least the same metric relations. (Coombs, 1964, p. 526-527)

Metric Models for Multidimensional Stimulus Generalization

There are three major points of view concerning rules of combination in multidimensional stimulus generalization. Two-dimensional models of these views are presented graphically in Fig. 2. The two stimulus dimensions are labelled S1 and S2, respectively, and their point of intersection represents the training stimulus. The amount of generalization is depicted as a third dimension, labelled R, orthogonal to the other two. The stimulus dimensions are assumed to be appropriately scaled in units of behavioral control so that the generalization gradients for the separate dimensions are linear and of equal slope. The task of the models is to predict the height of the generalization surface at any point in the four quadrants of the stimulus plane.

Theoretical justification for the first model, the excitation model, comes initially from Pavlov who proposed that generalization was due to a wave of excitation irradiating from the spot on the cortex stimulated as a result of the presentation of a stimulus:

> It may be assumed that each element of the receptor apparatus gains representation in the cortex of the hemispheres through its own proper central neurone, and the peripheral grouping of the receptor organs may be regarded as projecting itself in a definite grouping of nervous elements in the cortex. A nervous impulse reaching the cortex from a definite point from the peripheral receptor does not give rise to an excitation which is limited within the corresponding cortical element, but the excitation irradiates from its point of origin over the cortex, diminishing in intensity the further it spreads from the center of excitation. (Pavlov, 1927, p. 186)

A somewhat over-simplified interpretation of this view is that generalization between two stimuli is a function of the physical distance between their cortical representations. The important point, however, is that generalization is not referred to fixed sensory dimensions but occurs as a "spread of excitation" in all directions in the cortex. Stimuli are presumed to "acquire" the capacity to elicit a response because of certain electro-chemical effects which "spread out" from the initial point of stimulation.

Generalization was given a similar interpretation in Hull's neural-interaction theory. Hull (1943) postulated that generalization represented a "spread of habit strength" from the neural elements involved in stimulation by the conditioned stimulus to those activated by the test stimuli. This "spread of habit strength" from one physical stimulus to another was assumed to occur along innate afferent neural continua as an exponential decay function of the psychophysical similarity of the stimuli, as measured in discrimination units (j.n.d.'s).

These "spread of effect" theories imply that a generalization surface for two dimensions can be generated by rotating a unidimensional gradient around its peak, resulting in a surface which is conical in shape, as depicted in the first panel of Fig. 2. The most important psychological implication of this structure is that for a generalization surface of given dimensionality, one set of orthogonal dimensions, or one frame of reference, is as good as any other in predicting generalization behavior. Each pair of stimuli in the space defines a dimension along which generalization may occur.

According to the views of Pavlov and Hull, prior differential training along specific stimulus dimensions is not necessary to obtain gradients of generalization. A contrasting view, proposed by Lashley and Wade (1946), is that generalization is a performance phenomenon reflecting only an organism's failure to discriminate relevant aspects of a stimulus. They specifically reject Hull's claim that generalization represents a gradient of habit strength developed during conditioning. They argue that if an organism responds (generalizes) to a test stimulus it is only because it has not yet been conditioned to respond differentially to that stimulus difference along the relevant continuum, or it is not attending that aspect of the stimulus. If discrimination occurs, as reflected in generalization decrement, it occurs along stimulus dimensions previously established for the organism by differential training.

This view, according to Guttman (1956), predicts that generalization decrement for any stimulus varying in two dimensions is simply the sum of the decrements occurring along each of the component dimensions. This is the discrimination model depicted in the second panel of Fig. 2 and it is represented by a generalization surface that has the appearance of a pyramid, with its four corners on the stimulus axes. The significant feature of this structure is that it is determined by a fixed set of dimensions which are not subject to arbitrary rotation.

The *dominance* model is suggested by the results of the experiments, discussed earlier, in which some one stimulus property appeared to be "dominant" in exerting behavioral control. This model states that stimulus control is selective--exerted only by the dimension along which the most discriminable stimulus change occurred. Theoretical background for this model may be found in Lashley's "principle of dominant organization." Lashley (1942) proposed that the mechanism of nervous integration may be such that, when any complex of stimuli arouses nervous activity, that activity is immediately organized and certain stimulus properties become dominant for reaction while others become ineffective. The generalization surface predicted by the *dominance* model is depicted in the third panel of Fig. 2; it is also a pyramid with its corners at 45° to the stimulus axes.

If we look at successive sections through the generalization surfaces of Fig. 2, sections taken horizontal to the respective stimulus planes, sets of concentric contours are revealed. The contours are circular for the excitation model, diamond-shaped (rotated squares) for the discrimination model, and square for the dominance model. Each contour corresponds to a prescribed amount of generalization decrement and describes a locus of equally effective stimuli. That is, all stimuli on a prescribed contour are equivalent with respect to the amount of control exercised over differential responding, and thus are equally substitutable, one for another, to produce the prescribed generalization decrement. These are the level contours discussed in the preceeding section, whose shape uniquely determines the metric appropriate for a given set of measures. Thus, the assumption of the excitation model that all contours are circular is equivalent to the very special assumption that distances, in units of behavioral control, are Euclidean, and the combination rule predicting the generalization decrement to stimuli varying in two dimensions is the square root of the sum of squared decrements occurring on each component dimension. This is the combination rule adopted by most multidimensional scaling models.

The metric underlying the discrimination model provides an engagingly simple rule for combining distances along component dimensions. Namely, the distance between two points is equal to the sum of the differences between the projections of the two points on each of their coordinate axes. This is often called the "city-block" metric because the distance between two points is the total distance that must be traversed in a north-south direction, plus the distance that must be traversed along an east-west direction in order to get from one point to another. Theoretical arguments for this rule of combination have been proposed by Landahl (Householder and Landahl, 1945, p. 76; Landahl, 1945) and by Restle (1959).

The dominance metric requires that the distance between two points is simply the greatest of the distances separating their projections on each of the coordinate axes. Although this combination rule differs markedly from that of the "city-block" metric, in terms of its psychological implications and with regard to the question of how behavioral control is distributed among the component dimensions of a complex stimulus, mathematically the "city-block" space and the "dominance" space are isometric. Their isometry is evidenced by the observation that a 45° rotation of the stimulus plane (stretched by a factor $\sqrt{2}$) maps the level contours of the "city-block" metric into those of the "dominance" metric. Two points are served by this observation: (1) distances are not preserved under rotation of axes, that is, "city-block" distances are mapped into "dominance" distances, thus the dimensions of stimulus control for a given combination rule are unique; and, (2) unless the effective stimulus dimensions are known in advance of analysis, the combination rule generating the data cannot be uniquely determined. This latter point has implications for multidimensional analyses concerned simultaneously with determination of the effective dimensions of psychological space and its metric. (Shepard, 1962; Kruskal, 1964)

__A General Minkowski Model for Multidimensional Stimulus Generalization.__ Each of the three models considered for stimulus generalization provides differential predictions of the quantitative properties of behavior under complex stimulus control. It turns out that the metrics associated with these models are special instances of a single parameter family of distance functions known as the Minkowski r-metrics. For any $r \geq 1$ the r-distance between points $x = (x_i, \ldots, x_n)$ and $y = (y_1, \ldots, y_n)$ is defined to be

$$d_r(x,y) = \left[\sum_{i=1}^{n} \left| x_i - y_i \right|^r \right]^{1/r}$$

where the index i ranges over the n dimensions of the space and $|x_i - y_i|$ is the distance between the projections of x and y on the i^{th} dimension.

In application to two-dimensional stimulus generalization we equate the measure of generalization decrement from θ to stimulus $p = (p_1, p_2)$ with the distance $d_r(\theta , p)$, hence

$$\bar{g}(p) = d_r(\theta, p) = \left[x^r + y^r \right]^{1/r}$$

where $x = \bar{g}(p_1)$ and $y = \bar{g}(p_2)$.

In the (x,y) plane the level contours corresponding to $|x|^r + |y|^r = 1$ are called the "unit circles" (U_r) for different valuations of r. As shown in Fig. 3, U_2 is the Euclidean circle of radius 1; U_1 is an inscribed square with vertices $(\pm 1, 0)$ and $(0, \pm 1)$. As r increases to infinity, d_r approaches the distance

$$\lim_{r \to \infty} d_r = \max \{|x| , |y|\}$$

Thus, U_∞ is a circumscribed square with vertices $(\pm 1, \pm 1)$. As r increases from 1 to ∞, U_r deforms in a continuous manner from the square corresponding to d_1 to the square corresponding to d_∞. Intermediate values of r are represented by contours bounded by U_1 and U_∞. The "unit circle" for any fixed r contains those for smaller r; as shown in Fig. 3, $U_1 \subset U_2 \subset U_4 \subset U_\infty$.

The significance of r as a parameter of multidimensional stimulus control may be understood in connection with the question: How much does each component property of a compound stimulus contribute to the total behavioral control exerted by that stimulus? Or, more precisely: How are each of the effects that are attributed to the separate stimulus components weighted in the combination rule that predicts multidimensional stimulus control? If, for the compound stimulus $p = (p_1, \ldots , p_n)$, which we assume to have only positive components, we let x and x_i represent $\overline{g}(p)$ and $\overline{g}(p_i)$, respectively, then the Minkowski r-metric defined for $\overline{g}$ can be represented as a weighted sum of all component effects, namely

$$x = \Sigma\, \omega_i x_i$$

where $0 \le \omega_i \le 1$ $(i = 1, \ldots , n)$, $1 \le \Sigma\, \omega_i \le n$, and ω_i is the weight associated with x_i and dependent upon x_i in a manner determined by the parameter r.

If $r = 1$ (the discrimination model), then ω_i is a constant equal to 1 independent of x_i. All effects summate without differential weighting. On the other hand, if $r = \infty$ (the dominance model), then $\omega_i = 0$ for all x_i except $\max_i \{x_i\}$, for which $\omega_i = 1$. In this case behavior is interpreted as being controlled exclusively by the unidimensional stimulus change which, when occurring alone, produces the most generalization decrement.

In the case of the Euclidean combination rule, $r = 2$ (the excitation model),

$$x = \Sigma \cos \alpha_i\, x_i$$

that is, $\omega_i = x_i/x = \cos \alpha_i$ $(i = 1, \ldots , n)$ are the direction cosines of the p'-vector in psychological space, and α_i is the angle formed by the p'-vector and the p_i'-vector at their origin θ. This means that the weight assigned to each component effect x_i is proportional to the magnitude of x_i. Thus, large effects are proportionately weighted more than small effects in the prediction of $\overline{g}(p)$.

For arbitrary r, since $x^r = \Sigma\, x_i^r$, x can be written as

$$x = \Sigma\, (x_i/x)^{r-1} x_i \qquad \text{or} \qquad x = \Sigma\, \mathrm{cm}^{r-1} \alpha_i\, x_i,$$

where cm α_i denotes the Minkowski cosine of the angle α_i, pertaining to the relative position of the p_i'-vecotr to the p'-vector in Minkowski space (see Petty, 1955). (Definitions of trigonometric functions in Minkowski geometry are analogous to those of their Euclidean counterparts. The Minkowski cosine enters some of the same trigonometric identities as does the Euclidean cosine; however it cannot be interpreted as a function of a real number, "the angle formed by aribtrary vectors A and B", because its valuation depends on the order in which the vectors are considered. In general, $\mathrm{cm}(A,B) \neq \mathrm{cm}(B,A)$; however, if $\mathrm{cm}(A,B) = \mathrm{cm}(B,A)$ for all A, B in the vector space, then its geometry is Euclidean.)

As r increases, the separate unidimensional effects become increasingly disproportionately weighted by the combination rule that predicts multidimensional stimulus control (to the emphasis of relatively large effects and de-emphasis of relatively small effects). For example, if $0 < x_1 < x_2$, then their respective weights ω_1 and ω_2 are related in the following way:

$$\omega_1/\omega_2 = (\mathrm{cm}\, \alpha_1/\mathrm{cm}\, \alpha_2)^{r-1} = (x_1/x_2)^{r-1}$$

or

$$\omega_1 = (x_1/x_2)^{r-1}\, \omega_2.$$

Thus $\omega_1 < \omega_2$ for any $x_1 < x_2$ and $r > 1$, since $(x_1/x_2)^{r-1}$ is always less than 1 (when $r > 1$) and specifically tends to zero as r increases to $+\infty$. Consequently, as r increases, the weights become increasingly disproportionately distributed over the terms of the summation formula.

This formal result may serve the informal psychological notion that the extent to which an organism "attends" a given unidimensional stimulus change may depend on the changes that occur simultaneously along alternative dimensions of the stimulus. An organism is said to "attend" a particular dimension or property of a stimulus if its behavior is predominantly under the control of variations along that dimension. "Attention is a controlling *relation*--the relation between a response and a discriminative stimulus.... The criterion [of attending] is whether the stimulus is exerting any effect upon our behavior." (Skinner, 1953, p. 123-124) A given unidimensional stimulus change may be attended when occurring alone but not attended (or attended less) when occurring simultaneously with a more discriminable change along another dimension. In other words, the controlling relations between unidimensional stimulus changes and the response may alter when these changes occur in concert. The dominance metric applies to outcomes in which only a single dimension is attended, namely, that dimension along which the most discriminable change occurred. On the other hand, the city-block metric applies when these controlling relations remain constant, regardless of whether a given unidimensional change occurs alone or in the context of other changes. A Euclidean metrization may be interpreted simply as meaning that the organism attends separately each component dimension of a stimulus in proportion to the relative magnitude of the change in that dimension. By extension, intermediate forms of Minkowski r-metric may represent (by the weights ω_i) other ways in which attention may be distributed among the component properties of a compound stimulus.

Empirical Evidence of the Metric Properties of Multidimensional Stimulus Control

Butter (1963) conducted two operant conditioning experiments with pigeons (10 *Ss* in the first experiment and 27 *Ss* in the second), in order to compare generalization to stimuli varied in one and in two dimensions. In both experiments the pigeons were trained to peck at a circular translucent key upon which was projected a narrow (1/16") vertical strip of monochromatic (550 mμ) light. The birds were tested under extinction conditions for generalization to stimuli varied in the wavelength dimension, in the angular orientation (tilt) dimension, and in both dimensions. In the first experiment three levels of wavelength (520, 550, and 580 mμ) were crossed with three levels of angular orientation of the band of light (40, 90, and 140°), resulting in 9 different stimuli, plus a tenth consisting of a horizontal (0°) band of light (550 mμ). The stimuli were presented separately, 24 times each, for periods lasting 30 seconds. The second experiment was performed in the same manner with nine replications of each of 25 different stimuli, produced by crossing five levels of wavelength (530, 540, 550, 560, and 570 mμ) with five levels of angular orientation (30, 60, 90, 120, and 150°). In both experiments generalization to each wavelength-tilt combination was measured in terms of the average number of responses emitted furing the 30-second presentations of the stimulus.

Butter found that relative generalization along the separate dimensions of wavelength and tilt combined multiplicatively to predict relative generalization in both dimensions; that is, the obtained generalization measures entered the following relation:

$$n_{ij}/n_{oo} = (n_{io}/n_{oo}) \times (n_{oj}/n_{oo})$$

where n_{oo} denotes the number of responses emitted to the training stimulus, and n_{io}, n_{oj}, and n_{ij} denote, respectively, the number of responses emitted to a given change in wavelength, angular orientation, and both, respectively. This relation appeared to hold for the responses of individual Ss, as well as for group means in the first experiment. In the second experiment, predictions were less accurate for the results of individual Ss, but the relation held up very well for group means.

The contours that represent equal levels of generalization decrement in a graphic presentation of this combination rule have the shape shown in Fig. 1(c).

Although it was concluded that this finding is not consistent with the view that multidimensional generalization is the result of the algebraic summation of generalization decrements to all discriminable stimulus changes, a logarithmic transformation of the obtained generalization measures renders a simple additive relation between the corresponding measures of generalization decrement (hence, support for the discrimination hypothesis). The implication intended is that the response measures serving the definition of generalization in this experiment may enter a simple exponential relation with an alternative response variable, in terms of which multidimensional stimulus generalization has the simple metric properties implied by the discrimination model.

Blough (1965) has pointed out that response rate, that is, "the number of operants emitted in a fixed period of time," may be a poor choice of dependent variable for defining stimulus generalization because it is not a simple measure of response strength but, rather, is a composite measure containing a "number of responses to extinction" component, a "rapidity of response in the presence of the stimulus" component, a "latency-following-stimulus-onset" component, and a "response-probability, having-once-responded" component. How these separate components combine to determine the total number of operants emitted during an interval of time to a generalized stimulus is in no way clearly understood. The practice of averaging response rates over separate presentations of the stimulus (occurring at different times during extinction) and over separate subjects (for whom the response components may interact in different ways) makes the measure even more difficult to interpret simply. In view of these complexities there is no reason to prefer the initial form of this response variable, as a measure of generalization, to a logarithmic function of it--particularly since the latter has the metric propetties we desire.

Evidence for a simple additive rule of combination in terms of log response rate was also obtained in a generalization study cited earlier

(White, 1958). White trained three groups of 24 kindergarten children to pull a response handle freely during 3.8-sec presentations of a colored stimulus (Munsell patch: 10GY 8/6). Candy was used as reinforcement. One group was tested on the training stimulus and three novel stimuli of the same lightness (value) but having graded differences in hue (5GY, 10Y, and 5Y 8/6). A second group received test stimuli identical to the training stimulus in hue but differing in lightness (10GY 7/6, 6/6, and 5/6). The test stimuli administered to the third group differed from the training stimulus in both hue and lightness (5GY 7/6, 10GY 6/6, and 5Y 5/6). White reported that there was less generalization to stimuli differing in both dimensions than to those differing in either dimension alone. Converting the reported response measures to measures of relative generalization (response rate to training stimulus divided by response rate to test stimulus), it turns out that the generalization functions for the three groups are almost perfectly related multiplicatively, that is, relative generalization to a given change in hue and value equals the product of relative generalization to hue alone, times that to value alone. This relation is the same as that found by Butter.

Numerous studies have been undertaken in the area of multidimensional psychophysics in order to determine what dimensions or phychological attributes of complex stimuli control human judgments of stimulus similarity and dissimilarity, and to derive, from these judgments, scales for the attributes in question. By what principle subjects combine, in a single judgment, differences along several stimulus dimensions has, in general, not been the problem under investigation. Attneave (1950) was the first in this area to investigate directly the combination rule pertaining to multidimensional stimulus similarities. He questioned the appropriateness of the Euclidean assumption underlying multidimensional scaling procedures on the grounds that it implies either that perceptual judgments are not referred to a unique set of psychological dimensions or that a single judgment subjects combine differences along several stimulus dimensions, according to the square root of the sum of squared differences on each dimension. He viewed both alternatives as objectionable and argued, instead, for an alternative hypothesis which holds tht the dimensions of psychological space are unique, and that the perceived difference between two stimuli is equal to the simple arithmetic sum of their differences on the individual dimensions.

Attneave undertook one experiment in which 100 Ss were asked to rate on a seven-point scale the pairwise similarities of seven squares differing in size and/or brightness. The squares were cut from grey paper and mounted in pairs on black cardboard rectangles. Three squares were 1.5" on each side with reflectances of 5.3, 8.8, and 37.3 percent, respectively; three squares had a reflectance of 20.7 percent, with sides of 1.0, 2.0, and 2.5", respectively; and a seventh square had sides of 1.5" and a reflectance of 20.7 percent. Thus the stimuli formed a "plus" configuration in physical coordinates.

The ratings obtained for the similarity-dissimilarity of each of the 21 stimulus pairs were scaled by the Method of Graded Dichotomies (see Torgerson, 1958). To be considered as "distance", behavioral measures of stimulus differences along any single continuum must be linearly additive. Attneave, however, was unable to fix a zero-point for the judgment scale that would satisfy this condition for the stimulus dimensions involved. Rather, it was found that the judged differences between pairs of stimuli were linear with the differences between the logs of their corresponding physical values. Thus the appropriate psychological dimensions appeared to be the log (Fechner) transforms of the corresponding physical continua. In the face of this non-linearity a consistent metric treatment of similarities would have been virtually impossible had not the physical dimensions of the stimuli been known.

Judgments of the overall difference between stimuli differing in both dimensions were found to be related to the physical variables of area and reflectance by an equation of the following form:

$$X_{ij} = W_1 \log(A_i/A_j) + W_2 \log(R_i/R_j) + C$$

in which X_{ij} represents the scaled difference between stimuli i and j; C represents the additive constant, fixing the origin of the judgment scale; W_1 and W_2 are constants, optimally weighting the two dimensions; and A and R represent area and reflectance, respectively. The linear correlation between obtained measures and those predicted by the equation was 0.967.

In log-log coordinates of physical space the level contours corresponding to this outcome are affinely equivalent to the unit circles of the city-block metric.

Similar findings were obtained in two other experiments reported by Attneave: in one, parallelograms, differing in length of base and in angularity, with one also differing in color, served as stimuli; in another, triangles differing in size and in shape served as stimuli. Attneave concluded: "...the psychological differences between [stimuli] may be conceptualized as distances in a non-Euclidean space; this space has an axis system which is psychologically fixed, and hence not subject to rotation in the treatment of data; and a multidimensional distance in this space does not differ greatly from the sum of its projections on the axes." (1958, p. 551)

In the light of this outcome, and in view of the fact that the multidimensional scaling methods (despite their restriction to the Euclidean metric) have apparently achieved satisfactory solutions to data obtained with color stimuli, Torgerson (1958; p 254) and Shepard (1964; p 80) suggest that the form of the metric appropriate for a given set of data may depend upon the extent to which the stimuli differ with respect to "obvious and compelling" dimensions, or whether they are perceived as "homogeneous, unitary wholes." Torgerson (1958) pointed out:

> Attneave's choice of dimensions and of the stimuli along these dimensions presented very favorable conditions for obtaining the 'simple' sum type of judgment required by his model. The dimensions varied were simple, obvious, and compelling.... In addition, the stimulus pattern itself [a 'plus' in physical coordinates] would emphasize the 'given dimensions.' (p. 292)

Under these conditions a city-block metric might very well apply. On the other hand, when subjects are confronted with homogeneous patches of color a "different dimension of comparison is likely to be invoked for each pair of colors." (Shepard, 1964; p. 80) Under these conditions a Euclidean metric should be most appropriate.

The difficulty in this analysis lies in the fact that color stimuli do not always result in behavioral data that conform to a Euclidean metric. In White's experiment described above, for example, an additive metric was found to hold for generalization to Munsell color patches. This was also true in the Jones (1962) study. Jones showed a particular Munsell color (7.5PB 5/6) to 30 human Ss. Test colors were then presented, varying in saturation (chroma) only, brightness (value) only, hue only, or in all combinations of these properties, taken two or three at a time. Subjects were required to sort the test colors into compartments labelled +, 0, and ?, indicating, respectively, that they were the same as the standard color, different from it, or that S was uncertain whether they were the same or different. A score of 2 was arbitrarily assigned to the test color if it was sorted into the + compartment and a score of 1 if it was sorted into the ? compartment. Three presentations of each stimulus allowed a maximum possible score of 6. A measure of stimulus generalization was defined as the sum of scores assigned to each color averaged over all 30 Ss. Corresponding measures of relative generalization to colors differing from the standard stimulus in only one property were found to combine multiplicatively in predicting relative generalization to colors differing in two or more properties. A metric treatment of these results in terms of log transformations suggests tht Attneave's additive model applies to the color domain as well as to stimuli having "obvious and compelling" dimensions.

Shepard (1964) investigated the "match to sample" behavior of human Ss in an attempt to determine the shape of the isosimilarity contour and, thus, the particular Minkowski metric that determines how differences in size and tilt combine to control choices based on overall similarity. Schematic one-spoke wheels (circles with single radial lines) served as stimuli. The dimensions of variation were the diameter of the circle and the inclination of the radial spoke.

The standard stimulus was 3/4" in diameter with its spoke inclined 45° from horizontal. Twenty-four series of 15 different comparison stimuli were presented to each of 60 Ss with instructions to pick the one stimulus from each series that had the "closest overall resemblance" to the standard stimulus. In physically-defined space with coordinates of size and inclination the 15 stimuli in each series were uniformly spaced along straight lines tangential, at 24 different clockwise positions, to an imaginary circle with a center corresponding to the standard stimulus. The middle stimuli in the 24 series were all equidistant (in a physical sense) from the standard stimulus.

The distributions of frequency of choice were plotted separately for each of the 24 stimulus arrays. These distributions appeared to pass through four cycles of change as the test arrays rotated clockwise around the standard stimulus. At the four positions corresponding to unidimensional stimulus variation (constant size, varying tilt; or constant tilt, varying size) the distributions were sharply peaked at the middle stimulus of each array. As the arrays were rotated through positions at 45° angles to the physical axes (corresponding to simultaneous variation in both size and tilt) these distributions flattened out considerably; in fact, they appeared to become slightly bimodal. This result implied that the isosimilarity contour had four regions of sharpest curvature. Instad of conforming to a Euclidean ellipse, it presumably resembled a roughly "four-cornered" figure. In addition, the tendency toward bimodality of the distributions of choices along the four tangential series that ran at 45° through the physical representation implied that the "four-cornered contour must in fact have been concave.

Shepard suggested that this anomalous result was due to the indiscriminate pooling of data from subjects who were attending different aspects of the stimuli. The Ss were divided into two groups; the data from Ss whose behavior appeared to be under the control of both stimulus properties were analyzed separately from the data of those whose behavior appeared to be predominantly under the control of one or the other stimulus property. The Ss in the latter group were described as adopting a "matching strategy" which essentially involved "finding the best match along one of the two dimensions." Data from Ss of the other group, who were presumed to be attending both dimensions, indicated that the isosimilarity contour was convex, but still roughly "four-cornered", indicating the appropriateness of a Minkowski metric somewhere between the Euclidean and the city-block varieties.

In the second experiment, similarity was defined in terms of the proportion of times that two stimuli were confused during identification training. As in the preceding experiment each stimulus consisted of a circle of a certain size, together with a single radial line inclined at a certain angle. Only eight stimuli were used, however, and they were chosen so as to form the vertices of a regular octagon in the two-dimensional physical representation of the stimuli. Twenty Ss were required to learn a different arbitrarily specified letter of the alphabet as a label for each of the eight stimuli. The errors made during the learning of these stimulus-response assignments provided a "measure of similarity" for each pair of stimuli, based on the number of times either stimulus in the pair led to a response assigned to the other stimulus in that pair.

Shepard reasoned that "the assumption of a Euclidean metric would predict that about the same number of confusions should occur between any two stimuli that are represented by adjacent points around the octagon" because the two dimensions were scaled with respect to each other in such a way that equal displacements along either physical dimension would produce equivalent psychological changes. Even if some eccentricity remained in the isosimilarity contour the "number of errors between adjacent stimuli should pass through only two cycles as one revolution is completed about the octagon." (p. 74) On the other hand, the assumption of a city-block metric would predict that the number of confusions between the four pairs of adjacent stimuli that differed along only one dimension would always exceed the number of confusions between the four pairs of adjacent stimuli that differed along both dimensions at once. In this case the number of errors between adjacent stimulus pairs should pass through four cycles as one revolution is completed about the octagon.

The outcome was strongly in support of the second of these two predictions. Shepard found, however, that this outcome (based on averaged data) was not representative of the results that appeared to hold for individual Ss. A correlational analysis showed that Ss who frequently confused stimuli differing along one dimension on one side of the octagon also tended to confuse stimuli differing along that same dimension on the opposite side of the octagon, but confusions between stimuli differing along the other dimension were not correlated across sides. Approximately half of the Ss tended to confuse differences in size predominantly, and the other half confused differences in tilt. Thus, the metric suggested by the averaged data resulted from pooling the responses of Ss who were attending different properties of the stimulus.

The Problem of Attention. It has been found in several studies that different aspects of a stimulus may be important in controlling the behavior of different Ss, or of a single S during separate presentations of the same stimulus. Revesz (1925) trained monkeys in a 4-box Yerkes Multiple-Choice apparatus on a simultaneous discrimination in which responses to a yellow circle (S+) were reinforced and responses to a blue rectangle, a red triangle, or a green trapezium (S-), presented simultaneously with S+, were not reinforced. When the same four colors were presented on other forms no S responded to a stimulus that was not identical to S+ in either color or form. On some trials, however, it was found that a subject's response was under control of the form of the stimulus and on others it was under control of its color. Attention appeared to alternate between the two stimulus properties. Lashley (1938, 1942) showed that when rats form a simultaneous discrimination different aspects of the discriminative stimulus gain control over responding for different rats. This outcome was also found by Reynolds (1961). Reynolds trained two pigeons on a successive discrimination in which S+ was a white triangle on a red background and S- was a white circle on a green background. When the stimulus components were presented separately (i.e., white triangle alone, red background alone, etc.) one S responded only when the red background was presented, and the other S responded only when the white triangle was presented. Thus, only one aspect of S+ acquired control over responding, but it was a different one for each S.

These results lead us to question the extent to which attentional factors may have occurred in other experiments only to be disguised by the practice of averaging data. Quantitative analyses of stimulus control are invariably undertaken on average data, whether the averaging is performed over Ss, or over separate stimulus presentations for a single S, or both. It is true that replication is important for increasing the reliability of behavioral measures and that some measures are defined only on the basis of repeated stimulus presentations, but averaged data is likely to have characteristics not to be found in individual data and it is likely not to be representative of momentarily effective S-R relations. This does not mean that averaging is bad or that valid inferences cannot be drawn from results based on averaging, but it does mean that interpretation of these results must be consistent with the facts of attention. In other words, in interpreting results it must be allowed that measures may not represent the homogeneous effects of all stimulus properties operating simultaneously but rather the composite effect of pooling separate S-R occurrences.

Experimental Conclusions and Problems for Further Research. The findings of several quantitative studies of multidimensional stimulus control indicate that when the relevant dimensions of control are presumed to be known the effects of multidimensional stimulus variation are amenable to metric treatment. This outcome holds for stimulus control established through experimental conditioning by the restriction of reinforcement contingencies to specific stimulus-response relations, as well as for stimulus control established, in the case of human Ss, through verbal instructions that serve to tap the pre-experimental history of each S. Rather than proving the need for alternative metric representations for data collected under different conditions, however, the findings based on quantitative measures of stimulus control tend to converge on the general appropriateness of a simple additive rule of combination.

For conditioning theory, this outcome implies the validity of Guttman's discrimination hypothesis; for psychophysics, it supports Attneave's contention that the psychological reference-systems underlying the perception of similarity and differences among stimuli are unique.

Certain questions may be raised with regard to the above findings that pertain to the generality and interpretation of these conclusions.

(1) Does a metric representation of multidimensional stimulus control necessarily imply that all Ss are attending the same stimulus aspects? To

what extent does a metric solution based on average data depend on the condition that the properties of a compound stimulus are not selectively effective in controlling differential responding? Shepard (1964) implies that if the responses of different Ss are controlled by different aspects of the stimulus, then averaging over Ss may result in measures that violate the triangle inequality property of distance. Would a nonmetric combination rule apply under these circumstances or could it perhaps be the case that an additive combination rule is generally found for averaged data because the separate aspects of a multidimensional stimulus are selectively attended upon repeated presentation to different Ss?

(2) In terms of what response variables does an additive combination rule apply in multidimensional stimulus generalization? A metric treatment of the data from each of the generalization studies discussed above involved logarithmic transformations of the experimental response measures (average rate of emission of a free operant in two cases and what was essentially a 3-point category scale measure in the third case). It was argued that the multiplicative structure of the data implied the availability of an alternative response measure in terms of which an additive metric would have been obtained directly. Unless this variable can be identified the point is only argumentative and there is no reason to prefer the metric combination rule over the multiplicative rule as a general principle of multidimensional stimulus control.

(3) To what extent do the stimulus properties that acquire behavioral control in a given experiment depend upon the direction and magnitude of prevailing stimulus variation with respect to these properties? The question here is essentially whether the effective dimensions of stimulus control represent a generally fixed frame of reference for perception or do they arise in the experimental situation as a consequence of the investigators special selection of stimuli? Attneave (1950) raised this question in considering the possibility

that the metric solution he obtained for measures of similarity among visual stimuli may have occurred only because his stimuli were selected in accordance with an obvious dimensional organization.

Problems of individual differences in S-R relations, the relativity of these relations (context effects), and questions concerning the inter-relations among alternative response variables are not unique to experiments employing complex stimuli but also arise to confound interpretation of experimental results obtained with unidimensional stimuli. Thus, further research directed toward the explication of these problems in multidimensional stimulus control will not only aid in extending the principles of stimulus control to the multidimensional case but may also provide illumination of problems generally encountered in behavioral analyses.

References

Attneave, F. Dimensions of similarity. *Amer. J. Psychol.*, 1950, 63, 516-556.

Blough, D. S. Definitions and measurement in general research. *Stimulus Generalization* (Mostofsky, D.I., editor), Stanford Univ. Press, 1965.

Brown, J. S. A linguistic analysis of generalization and discrimination. *Stimulus Generalization* (Mostofsky, D.I., editor), Stanford Univ. Press, 1965.

Busemann, H. The foundations of Minkowskian geometry. *Comm. Math. Helv.*, 1950, 24, 156-186.

Busemann, H. *The Geometry of Geodesics*. New York: Academic Press, 1955.

Butter, C. M. SLimulus generalization along one and two dimensions in pigeons. *J. Exp. Psychol.*, 1963, 65, 339-346.

Coombs, C. H. *A Theory of Data*. New York: John Wiley & Sons, 1964.

Eninger, M. U. Habit summation in a selective learning problem. *J. Comp. Physiol. Psychol.*, 1952, 45, 604-608.

Fink, J. B. & Patton, R. M. Decrement of a learned drinking response accompanying changes in several stimulus characteristics. *J. Comp. Physiol. Psychol.*, 1953, 46, 23-27.

Galanter, E. An axiomatic and experimental study of sensory order and measure. *Psychol. Rev.*, 1956, 63, 16-28.

Gibson, E. J. A re-examination of generalization. *Psychol. Rev.*, 1959, 66, 340-342.

Guttman, N. The pigeon and the spectrum and other perplexities. *Psychol. Rep.*, 1956, 2, 449-460.

Hancock, H. *Development of the Minkowsky Geometry of Numbers*. Vol. 1. New York: Dover, 1964.

Hanson, H. M. The effects of discrimination training on stimulus generalization. Unpublished doctoral dissertation, Duke Univ., 1956.

Harlow, H. F. Studies in discrimination learning by monkeys: VI. Discrimination between stimuli differing in both color and form, only in color, and only in form. J. Gen. Psychol., 1945, 33, 225-265.

Householder, A. S. and Landahl, H. D. Mathematical Biophysics of the Central Nervous System. Bloomington, Ind: Principia Press, 1945.

Hull, C. L. Principles of Behavior. New York: Appleton-Century Crofts, 1943

Jones, J. E. Stimulus generalization in two and three dimensions. Canad. J. Psychol., 1962, 16, 23-36.

Kruskal, J. B. Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. Psychometrika, 1964, 29, 1-27.

Landahl, H. D. Neural mechanisms for the concepts of difference and similarity. Bull. Math. Biophysics, 1945, 7, 83-88.

Lashley, K. S. An examination of the "continuity theory" as applied to discriminative learning. J. Gen. Psychol., 1942, 26, 241-265.

Lashley, K. S. The mechanism of vision: XV. Preliminary studies of the rat's capacity for detail vision. J. Gen. Psychol., 1938, 18, 123-193.

Lashley, K. S. & Wade, M. The Pavolvian theory of generalization. Psychol. Rev., 1946, 53, 72-87.

Mednick, S. A., and Freedman, J. L. Stimulus generalization. Psychol. Bull., 1960, 57, 169-200.

Messick, S. J., and Abelson, R. R. The additive constant problem in multidimensional scaling. Psychometrika, 1956, 21, 1-15.

Miller, J. The rate of conditioning of human subjects to single and multiple conditioned stimuli. J. Gen. Psychol., 1932, 20, 399-408.

Pavlov, I. P. Conditioned Reflexes, translated by G. V. Anrep. London: Oxford Univ. Press, 1927.

Petty, C. M. On the geometric of the Minkowski plane. Rivista Mat. Univ. Parma, 1955, 6, 269-292.

Prokasy, W. F., and Hall, J. F. Primary stimulus generalization. *Psychol. Rev.*, 1963, 70, 310-322.

Restle, F. A metric and an ordering on sets. *Psychometrika*, 1959, 24, 207-220.

Revesz, G. Experimental study in abstraction in monkeys. *J. Comp. Psychol.*, 1925, 5, 293-343.

Reynolds, G. S. Attention in the pigeon. *J. Exp. Anal. Behav.*, 1961, 4, 203-208.

Schoenfeld, W. N., and Cumming, W. W. Behavior and perception. *Psychology: A Study of a Science*, Vol. 5 (Koch, S., editor). New York: McGraw-Hill, 1963.

Shepard, R. N. The analysis of proximities I. *Psychometrika*, 1962, 27, 125-140.

Shepard, R. N. Attention and the metric structure of the stimulus space. *J. Math. Psych.*, 1964, 1, 54-87.

Shepard, R. N. Stimulus and response generalization: tests of a model relating generalization to distance in psychological space. *J. Exp. Psychol.*, 1958, 55, 509-523.

Skinner, B. F. *Science and Human Behavior*. New York: MacMillan, 1953.

Torgerson, W. S. Multidimensional scaling: I. Theory and method. *Psychometrika*, 1952, 19, 401-419.

Torgerson, W. S. *Theory and Methods of Scaling*. New York: Wiley, 1958.

Warren, J. M. Additivity of cues in a visual pattern discrimination by monkeys. *J. Comp. Physiol. Psychol.*, 1953, 46, 484-486.

Warren, J. M. Pe-ceptual dominance in discrimination learning by monkeys. *J. Comp. Physiol. Psychol.*, 1954, 47, 290-293.

White, S. W. Generalization of an instrumental response with variations in two attributes of the CS. _J. Exp. Psychol._, 1958, 56, 339-343.

Young, G., and Householder, A. S. Discussion of a set of points in terms of their mutual distances. _Psychometrika_, 1938, 3, 19-22.

Figure Captions

Fig. 1. Level contours arising from nonlinear stimulus-response relations.

Fig. 2. Possible forms of the generalization surface in two dimensions.

Fig. 3. Euclidean graphs of Minkowski "unit circles."

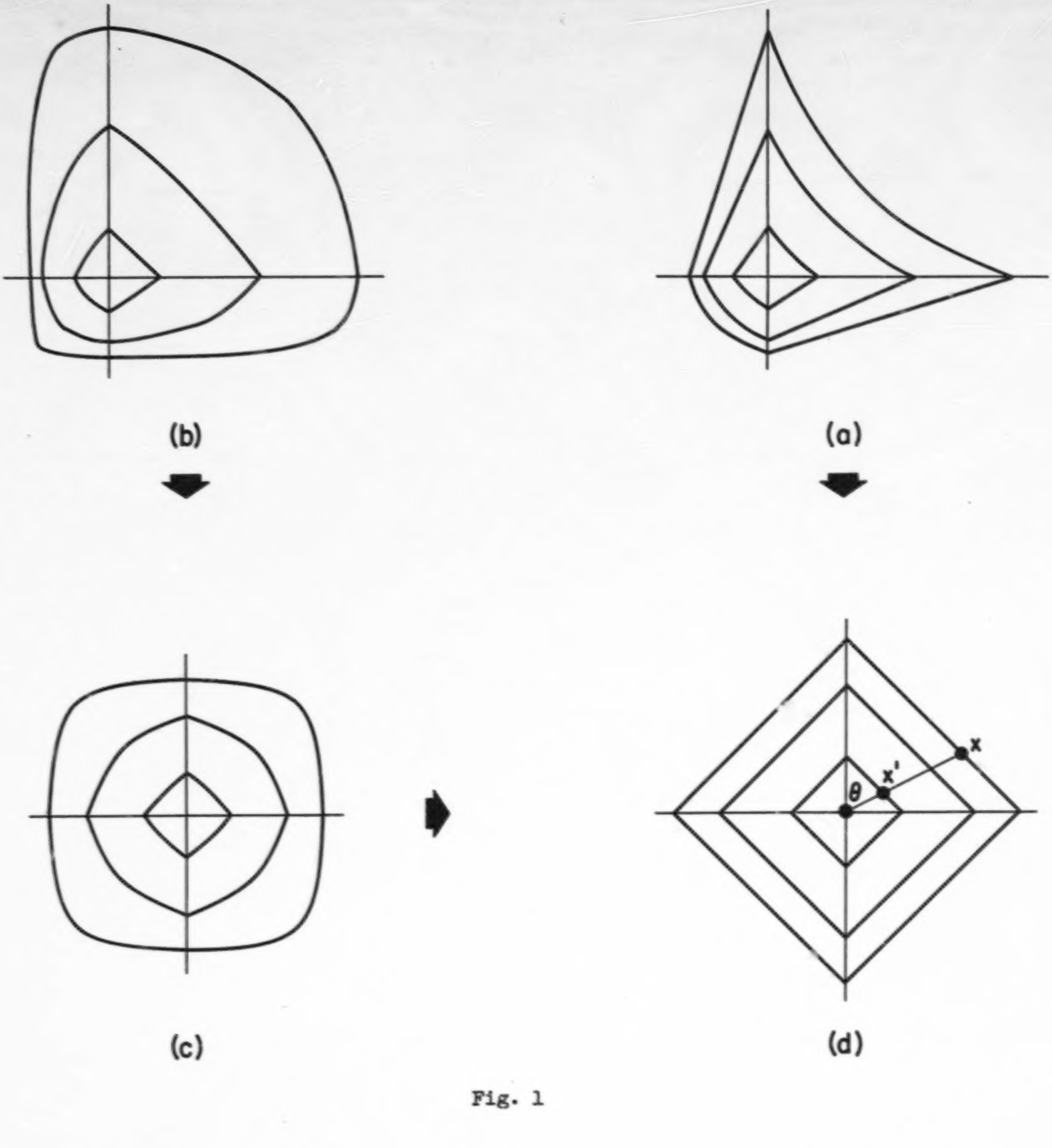

Fig. 1

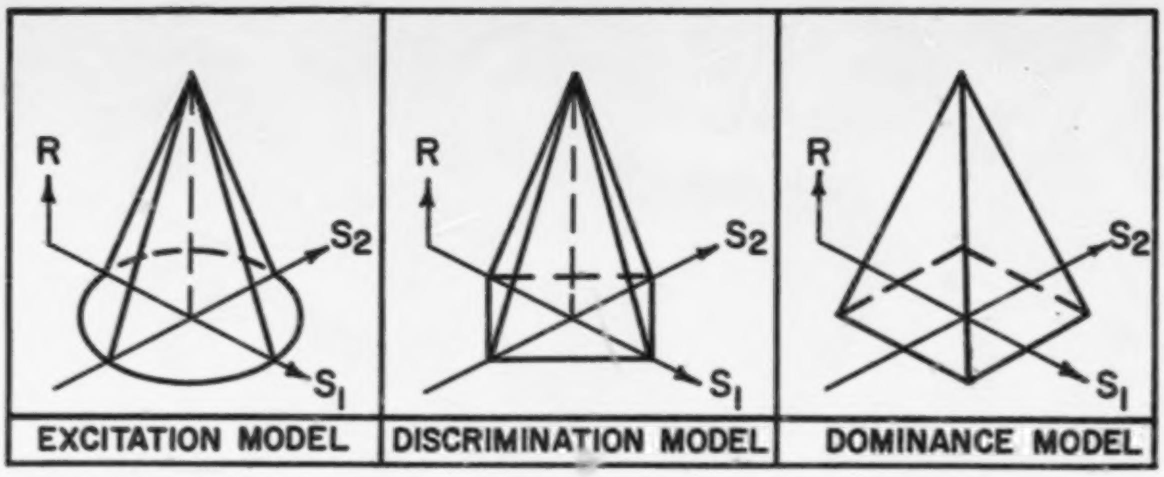
R
S2
S1
EXCITATION MODEL
R
S2
S1
DISCRIMINATION MODEL
R
S2
S1
DOMINANCE MODEL

Fig. 2

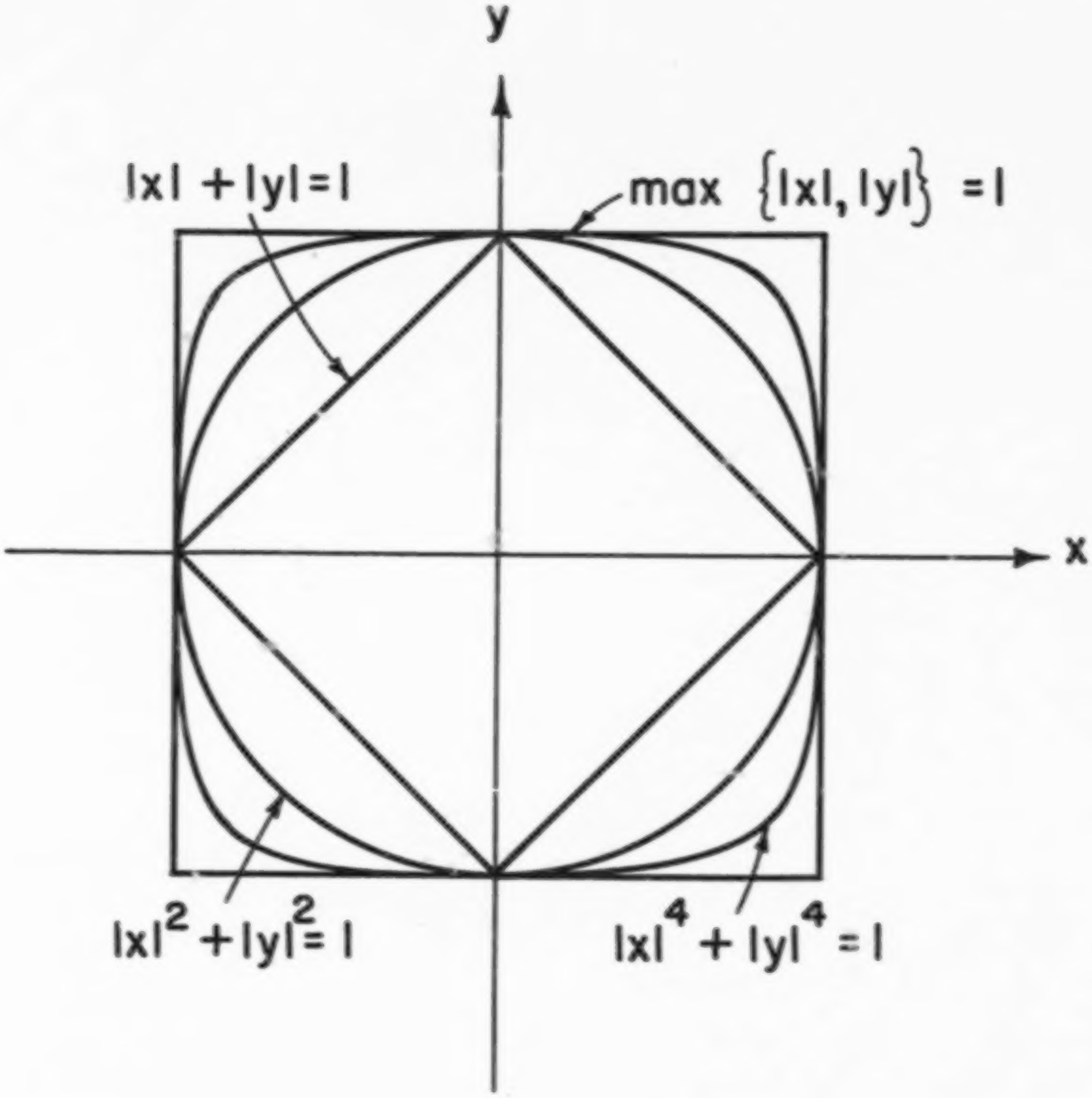
y
x
|x| + |y| = 1
max {|x|, |y|} = 1
|x|² + |y|² = 1
|x|⁴ + |y|⁴ = 1

Fig. 3